**BAB II**

**LANDASAN TEORI**

* 1. **Tinjauan Studi**

Adapun penelitian yang terkait tentang Clustering dan penggunaan metode *K-Means*, seperti di bawah ini :

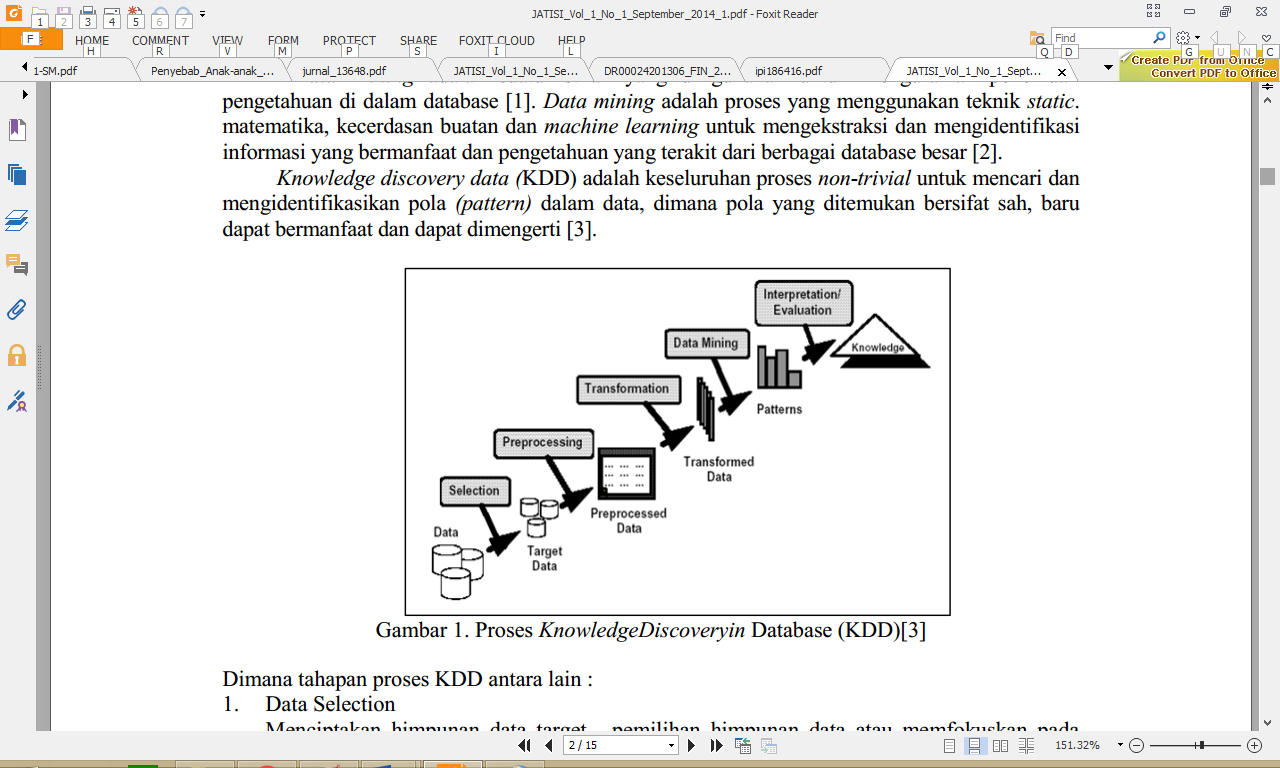
**Tabel 2.1** Penelitian Tekait

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Peneliti | Metode | Judul | Keterangan |
| 1 | *Yani Sugiyani.2016*[1] | *K Means* | Pengelompokan wilayah bedasarkan Potensi Hasil Pertanian Menggunakan Algoritma K-means Dikota Cilegon[1] | Dari hasil analisa pada penelitian tersebut bertujuan untuk mempermudah pemerintah memperoleh informasi data pengelompokan wilayah,  Dan akan menjadi bahan pengambilan kebijakan dalam hal peningkatan hasil pertanian. |
| 2 | Lianna Felicia, 2015[2] | *K-Means* | “Penerapan metode k-means clustering untuk memetakan potensi tanaman padi di kota Semarang” | Dari hasil penerapan k-means unntuk memetakan potensi tanaman dapat mempermudah pengelompokan hasil padi terbesar,sedang dan rendah. |

* 1. **Tinjauan Pustaka**
     1. **Data Mining.**

Data mining merupakan serangkaian proses penambangan data berupa informasi yang selama ini tidak diketahui secara manual dari suatu basisdata, Data mining juga digunakan untuk mencari ilmu pengetahuan yang bernilai tinggi dalam basis data yang jumlahnya besar sehingga disebut *Knowledge discovery in databases* (KDD).

Tahapan KDD sebagai berikut :



**Gambar 2.1** Tahapan *Knowledge Discovery in Databases*

“Data mining dapat diklasifikasikan dalam dua kategori yaitu deskriptif dan prediktif. mengklasifikasi sifat umum suatu data didalam databases disebut deskriptif sedangkan untuk mengambil kesimpulan terhadap data terakhir untuk membuat prediksi disebut prediktif”[4]

Berikut ini adalah beberapa teknik dalam data mining :

1. *Classification (Prediktive)*
2. *Clustering (Descriptive)*
3. *AssociationRule Discovery (Descriptive)*
4. *Regression (Prediktive)*
5. *DeviationDetection (Prediktive)*

Hasil dari analisa data mining sering dijadikan rekomendasi dalam pengambilan keputusan. Seperti pada contoh hasil analisa data mining berupa aplikasi bisnis informasi dapat digunakan daalam menejemen pemasaran produk agar strategi pemasaran yang dilakukan dapat diuji. Penggabungan antara data mining dan decision support system ini memerlukan evaluasi yang menjamin bahwa hanya hasil yang terbaik yang dapat digunakan.

Salah satu fungsi dari *postprocessing* atau evaluasi adalah memvisualisasikan dan memungkinkan analisis untuk mengeksplor hasil dan data *data mining* dari berbagai perspektif. Ukuran-ukuran statistik dan metode pengujian hipotesis dapat digunakan selama postprocessing untuk membuang hasil *data mining* yang tidak layak atau palsu.

Pada **Gambar 2.2** menunjukan hubungan antara data mining dengan area area lain.

Data mining juga secara khusus munggunakan ide-ide seperti : (1) estimasi,pengambilan contoh, dan pengujian hipotesis dari statistika dan (2) teknik pemodelan, algoritma pencarian, dan teori pembelajaran dari machine learning, pengenalan pola, dan kecerdasan buatan.



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

Sumber : “[https://www.docsity.com/en/data-mining-pertemuan-ke-satu-mahasiswa-stikom-tunas-bangsa/2175018”/](https://www.docsity.com/en/data-mining-pertemuan-ke-satu-mahasiswa-stikom-tunas-bangsa/2175018) [6]

* + 1. ***Clustering***

Metode clustering bertujuan untuk mengelompokan data yang mempunyai kemiripan yang sama antara satu data dengan data lain dan bersifat tanpa arahan atau diterapkan tanpa ada latihan, tidak memiliki target output dan tanpa guru.Dan memiliki dua jenis metode pengelompokan data yaitu *hierarchical clustering* dan *non-hierarchical clustering* [7]

Metode clustering memiliki potensi mengetahui struktur dalam data yang dapat gunakan selanjutnya untuk berbagai macam aplikasi secara luas seperti pengenalan pola,klasifikasi, pengolahan gambar [8]

*“Hierarchical clustering* merupakan metode yang mengelompokan dua atau lebih objek yang memiliki kesamaan paling dekat dan diproses ke objek yang lain dan seterusnya hingga membentuk tingkatan yang jelas antara objek dari yang tidak mirip sampai yang paling mirip, seadangkan metode *non-hierarchical*  
 menentukan terlebih dahulu jumlah cluster yang diinginkan dua tiga sampai seterusnya, Kemudian dilakukan proses cluster tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering”*[9]

* + 1. **Algoritma *K-Means.***

*K-Means* adalah data salah satu metode pengelompokan nonhierarki (sekatan).Dimana data dipartisi menjadi dua atau lebih kelompok sehingga data yang sama dimasukkan ke dalam satu kelompok yang sama dan data yang berbeda dikelompokkan kedalam kelompok yang lain.

Berikut tahapan melakukan clustering:

1. Menentukan jumlah *cluster k* yang di inginkan.
2. Pusat *cluster* ini dilakukan dengan cara random.
3. Menentukan pusat cluster berdasarkan jarak kedua objek tersebut.
4. Menghitung jarak menggunakan teori eucliaden dengan rumus berikut :

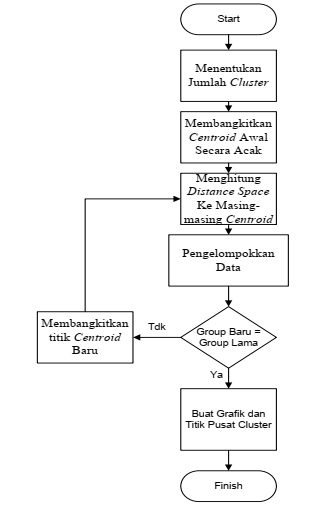


Diketahui:  
*D (i,j)* = Jarak data ke *i* ke pusat cluster *j*

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

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

1. Menghitung cluster baru.
2. Setelah memakai cluster baru jika hasil cluster tidak berubah maka perhitungan selesai dan jika hasil cluter berbeda maka proses cluster belum berhenti sampai hasil cluster sama dengan perhitungan sebelumnya



**Gambar 2.3** Tahapan Algoritma *K-Means*

* + - * 1. **Contoh Penerapan Metode K-Means Untuk Clustering**

Berikut ini merupakan contoh perhitungan K-Means Untuk Clustering Menentukan Barang yang Laris Terjual. Adapun langkah penyelesaiannya adalah sebagai berikut (Varuna Dewi, 2018).

Berikut tabel data yang dilakukan percobaan perhitungan manual :

**Tabel 2.2 Data Penjualan yang akan di Hitung**

|  |  |  |
| --- | --- | --- |
| **Nama Barang** | **Harga Barang (Rp1000)** | **Total** |
| Gordyn Black out | 130 | 600 |
| Gordyn Standart | 60 | 700 |
| Gordyn Beludru | 120 | 570 |
| Gordyb Silk | 110 | 590 |
| Vitrase Tile | 85 | 300 |
| Vitrase Organdi | 115 | 650 |
| Vitrase Turkie | 120 | 550 |
| Blind Roler | 130 | 585 |
| Blind Vetrikal | 140 | 70 |
| Blind Slim | 125 | 90 |
| Carpet Crown | 400 | 79 |
| Carpet Sandrio | 160 | 120 |
| Carpet Treasure | 950 | 105 |
| Carpet Spontini | 1750 | 30 |

**Iterasi ke-I**

1. **Menentukan Jumlah Cluster**

Tahap awal dalam proses *clustering* adalah menentukan berapa jumlah *cluster* yang diinginkan. Pada sistem Menentukan Barang yang Laris Terjual akan menggunakan dua titik cluster C1 dan C2

1. **Menentukan Pusat *Cluster (centroid)***

Penentuan pusat cluster pertama dilakukan secara random sehingga didapatkan sebagai berikut :

C 1: **(52. 28)**

C 2: **(50. 30)**

1. **Hitung Jarak Data dengan *Euclidean Distance.***

Euclidian distance digunakan untuk menghitung antara jarak data dengan pusat cluster dengan rumus sebagai berikut :



Dimana: x = Data.

y = Pusat cluster atau Centroid.

M1= (130,600) M6= (115,650) M11= (400, 79)

M2= (60,700) M7=(120,550) M12= (160,120)

M3= (120,570) M8= (130,585) M13= (950,105)

M4= (110,590) M9= (140,70) M14= (1750,30)

M5= (85,300) M10= (125,90)

Menghitung jarak dari semua data ketiap titik pusat awal:





























Kemudian Hitung jarak data ke pusat cluster atau centroid 2 :





























Dari perhitungan diatas dapat di hasilkan:

Hasil Iterasi Ke-1

|  |  |  |
| --- | --- | --- |
|  | **C1** | **C2** |
| **M1** | **577,29369** | **575,587** |
| **M2** | **672,04762** | **670,075** |
| **M3** | **601,8538** | **544,518** |
| **M4** | **620,71572** | **563,205** |
| **M5** | **273,99453** | **272,259** |
| **M6** | **625,18237** | **623,398** |
| **M7** | **526,41049** | **524,69** |
| **M8** | **562,43489** | **560,736** |
| **M9** | **97,508974** | **98,4886** |
| **M10** | **95,77578** | **96,0469** |
| **M11** | **351,71722** | **353,413** |
| **M12** | **141,87318** | **142,127** |
| **M13** | **901,29518** | **903,12** |
| **M14** | **1698,0012** | **1700** |

**C1 = M9,M10,M11,M12,M13,M14**

**C2 = M1,M2,M3,M4,M5,M6,M7,M8**

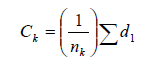
**Iterasi ke-II**

1. **Menentukan Jumlah Cluster**

Tahap awal dalam proses *clustering* adalah menentukan berapa jumlah *cluster* yang diinginkan. Pada sistem Menentukan Barang yang Laris Terjual akan menggunakan dua titik cluster C1 dan C2

1. **Menentukan Pusat *Cluster (centroid)***

Menetukan pusat centroid baru (Ck)



Dimana: nk = jumlah dokumen dalam cluster k.

d1 = adalah dokumen dalam cluster k.

Pusat cluster baru sebagai berikut:

**C1= (587,5 , 82,3)**

**C2= (108,75 , 568,1)**

1. **Hitung Jarak Data dengan *Euclidean Distance.***





























Selanjutnya hitung jarak data ke pusat cluster baru ke 2 yaitu :



























Pada hasil iterasi ke dua berubah maka proses berhenti

|  |  |  |
| --- | --- | --- |
|  | **C1** | **C2** |
| **M1** | **690,88316** | **38,309** |
| **M2** | **812,28661** | **307,716** |
| **M3** | **675,57941** | **11,4052** |
| **M4** | **849,25352** | **51,8901** |
| **M5** | **618,5744** | **286,254** |
| **M6** | **738,60649** | **82,1132** |
| **M7** | **661,28628** | **21,3326** |
| **M8** | **679,71578** | **27,1354** |
| **M9** | **447,84991** | **469,167** |
| **M10** | **468,27251** | **559,361** |
| **M11** | **187,52904** | **569,271** |
| **M12** | **429,15911** | **451,046** |
| **M13** | **363,21005** | **960,305** |
| **M14** | **1163,6759** | **1727,22** |

C1 = M9,M10,M11,M12,M13,M13,M14

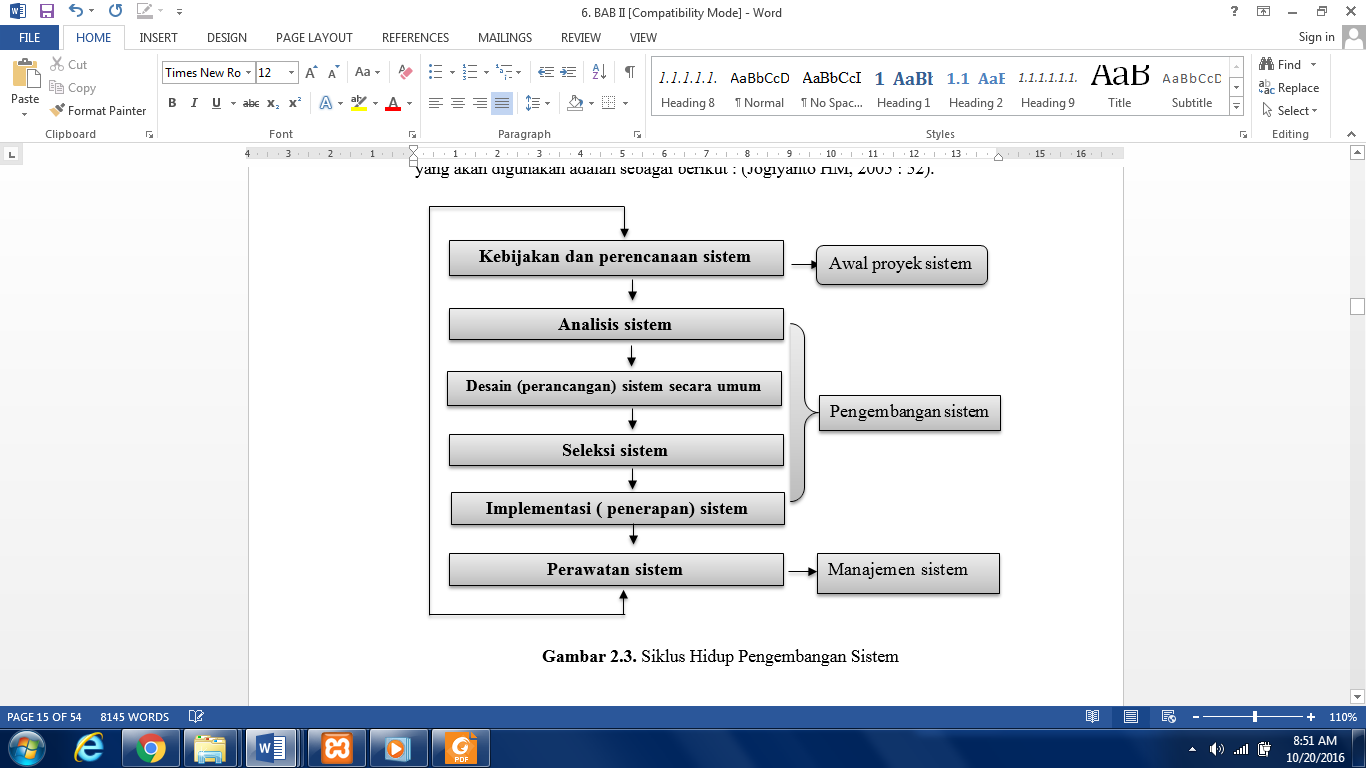
C2 = M1,M2,M3,M4,M5,M6,M7,M8

* + 1. **Siklus Hidup Pengembangan Sistem**

Dalam penyelesaian pengembangan system membutuhkan waktu yang lama sampai bertahun tahun diakbatkan karena proses pengembangan sistem dari system itu direncanakan sampai diterapkan,dipromosikan dan dipelihara. Ketika system yang telah dibuat mendapatkan kritikan atau masalah dan tidak bisa diatasi pada pemeliharaan system maka dilakukan perancangan system kembali.

*systems life cycle* atau siklus hidup dari pengembangan sistem adalah suatu bentuk yang diambil untuk menggambarkan tahapan utama dan langkah-langkah didalam tahapan tersebut dalam proses pengembangan.

Berikut langkah langkah utama yang digunaka pada Siklus hidup pengmbangan system yaitu :



**Gambar 2.4** Siklus hidup pengembangan sistem

Sumber : Jogiyanto HM, (2005 : 52 )[10]

Dalam *systems life cycle*, semua bagian dari pengembangan sistem dipisahkan menjadi berbagai macam tahapan kerja. Dan setiap tahapan memiliki karakteristik masing masing. Ada beberapa tahapan pengmbangan system sistem yaitu dari tahapan perencanaan sistem, analisis sistem, desain sistem, seleksi sistem, implementasi sistem dan perawatan sistem.

* + 1. **Perencanaan Sistem**

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

Hariyanto (2004 : 353)[11] 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 implementasi sistem”.

Perencanaan atau *planning* adalah hal-hal yang menyangkut studi tentang keperluan user dan studi kelayakan baik secara teknologi maupun secara teknis dan penjadwalan pengmbangan proyek perangkat lunak atau system informasi. Dimana pada tahapan perencanaan ini pengembang melakukan observasi untuk mengenali user baru dari sistem informasi atau perangkat lunak yang dikembangkan. DAD digunakan pada pengembangan system sebagai kakas (*tool*), semua permasalahan dimodelkan sebagai *use case* untuk mengvisualkan semua kebutuhan user atau pengguna,

* + 1. **Analisa Sistem**

Analisa sistem dapat diartika sebagai penjelasan dari suatu sistem informasi dimana bagian-bagian komponenya di identifikasi dan dievakuasi permasalahan atau hambatan dan kebutuhan yang diharapkan sehingga dapat diajukan perbaikannya. Dalam analisa system dapat mempelajari masalah dan kebutuhan sebuah organisasi untuk menentukan bagaimana orang, data, proses dan teknologi informasi dapat mencapai kemajuan terbaik untuk bisnis.

Whitten, et al. (2004 :33)[12] mengungkapkan “ *System analysis* adalah study domain masalah bisnis untuk merekomendasikan perbaikan dan menspesifikasi persyaratan dan prioritas bisnis untuk solusi”.

Dalam dunia analisis dan desain sistem Impak teknologi objek sangat diperlukan.. 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. sehingga, metode analisis dan desain berorientasi objek telah muncul sebagai pendekatan terpilih untuk membangun kebanyakan sistem informasi sekarang.

Dalam menyelesaikan pekerjaan seorang analis harus memiliki keahlian lain berupa (1) Pengalaman dan keahlian pemrograman computer dan (2) Pengetahuan umum proses dan teknologi bisnis:

Tahap analisis merupakan tahap yang penting dan kritis, jika terjadi kesalahan pada tahapan ini akan menyebankan kesalahan pada tahapan berikutnya. Berikut tahapan analisis :

1. Studi Kelayakan.

Untuk menetukan dimana berhasilnya solusi yang di ajukan terdapat pada tahapan study kelayakan.Dimana tahapan ini berguna untuk memastikan bahwa solusi yang diajukan benar-benar dapat dicapai dengan sumber daya dan dengan memperhatikan kendala yang terdapat pada perusahan serta dampak terhadap lingkungan sekeliling. Berikut tugas-tugas yang ada dalam study kelayakan meliputi :

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

Ada pun tugas-tugas lain sebagai berikut :

1. Mengusulkan 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 )”[13].

1. Analisis kebutuhan.

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

Berikut langkah-langkah yang harus dilakukan oleh analisis system yaitu :

1. Mengidentifikasi masalah.
2. Memahami kerja dari sistem yang ada.
3. Menganalisis sistem tanpa report..
4. Membuat laporan hasil analisis.

Tujuanya dalam pembuatan laporan hasil dilakukan ;

1. Pelaporan bahwa analisi telah selesai dilaksanakan.
2. Meluruskan kesalah pahaman tentang apa yang telah ditemukan dan dianalisis oleh analis sistem tetapi tidak sesuai menurut manajemen.
   * 1. **Desain Sistem**

Desain system merupakan tahapan yang dilakukan ketika analisis system telah dilakukan dimana pada tahapan analisis sudah mendapat gambaran apa yang harus dikerjakan.Desain sistem adalah spesifikasi atau intruksi solusi yang teknis dan berbasis komputer untuk persyaratan bisnis yang diidentifkasikaan dalam analisis sistem.

Tujuan utaman dalam tahan desain system adalah sebagai berikut :

1. Kebutuhan user atau pengguna dapat terpenuhi.
2. Menjadikan edukasi atau gambaran untuk pemogram komputer atau pihak lainya .

Perancangan sistem merupakan suatu keinginan membuat desain teknis yang berdasarkan evaluasi yang dilakukan pada kegiatan analisis dan bertujuan menghasilkan sua tu sistem komputerisasi.

Tedapat dua jenis perancangan system yaitu perancangan konseptual dan perancangan fifik :

1. P erancangan konseptual.

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 Romney, et al. 1997 dalam abdul kadir (2003 :407 )[13] “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)”.
   * 1. **Desain sistem secara umum**

Desain secara umum dilakukan oleh analisis sistem untuk mengidentifikasikan komponen-komponen sistem informasi yang akan didesain secara rinci oleh pemograman komputer dan ahli teknik lainya 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..

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

* + 1. **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 2.3** Bagan Alir Sistem

|  |  |  |  |
| --- | --- | --- | --- |
|  | **NAMA SIMBOL** | **SIMBOL** | **KETERANGAN** |
| 1. | Dokumen |  | Menunjukkan dokumen *input* dan *output* baik itu proses manual, mekanik, atau komputer |
| 2. | Kegiatan Manual |  | Menunjukan pekerjaan manual |
| 3. | Simpanan Offline |  | Menunjukkan file non-komputer yang diarsip urut angka (*numerical*), huruf (*alphabetical*), atau tanggal (*chronological*) |
| 4. | Kartu Plong |  | Menunjukkan *input* dan *output* yang menggunakan kartu plong (*punched card*). |
| 5. | Proses |  | Menunjukkankegiatan proses dari operasi program komputer |
| 6. | Operasi Luar |  | Menunjukkan operasi yang dilakukan di luar proses operasi komputer |
| 7. | Hard Disk |  | Menunjukkan *input* dan *output*  menggunakan *harddisk* |
| 10. | Keyboard |  | Menunjukkan *input* yang menggunakan *on-line keyboard* |
| 12. | Hubungan Komunikasi |  | Menunjukkan proses transmisi data melalui *channel* komunikasi. |
| 13. | Garis Alir |  | Menunjukkan arus dari proses |
| 14. | Penjelasan |  | Menunjukkan penjelasan dari suatu proses |
| 15. | Simbol Penghubung |  | Menunjukkan penghubung ke halaman yang masih sama atau ke halaman yang lain |

Sumber : Jogyanto, 2005 : 802[10]

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 2.5** 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 2.6** 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 2.7** Notasi Proses di DAD

1. *Data store* (simpanan data).

”Simpanan data pada DFD dapat disimbolkan dengan sepasang garis horisontal paralel yang tertutup disalah satu ujungnya”[10].

Media Nama Data store

**Gambar 2.8** Notasi Simpanan Data di DAD

* + 1. **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.

* + 1. **Pemeliharaan Sistem**

“Pemeliharaan sistem adalah proses pengubahan sistem setelah beroperasidan digunakan” [11].

“Pemeliharaan sistem adalah tahap dimana kita mulai pengoperasian sistemdan, jika diperlukan, melakukan perbaikan-perbaikan kecil” [14].

* + 1. **Teknik Pengujian Sistem**

“Tahapan akhir dari proses pengembangan perangkat lunak adalah pengujian. Menurut Pressman, pengujian perangkat lunak merupakan salah satu elemen dari rekayasa perangkat lunak yang sering disebut dengan verification and validation testing (V&V)” [15].

“Verifikasi sendiri mengacu pada serangkaian kegiatan yang memastikan perangkat lunak dapat melakukan fungsi-fungsi tertentu yang telah ditentukan. Validasi mengacu pada serangkaian aktivitas yang berbeda yang memastikan perangkat lunak sudah sesuai dengan kebutuhan pengguna. Tujuan dari pengujian perangkat lunak ialah untuk meningkatkan kualitas perangkat lunak, melakukan uji verifikasi dan validasi perangkat lunak yang dikembangkan, serta menguji reliabilitas perangkat lunak” [15].

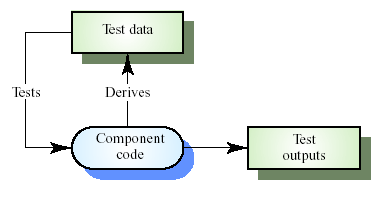
**2.2.12.1 *White Box***

Ada dua macam pengujian perangkat lunak yang dapat dilakukan, yaitu metode *white box* dan *blackbox. White box testing*  dilakukan di awal program, sedangkan *black box testing* baru dilakukan pada tahap berikutnya. Pengujian dilakukan untuk melakukan penilaian tentang keberhasilan program dan menentukan apakah prgram sudah sesuai dengan keinginan *user* atau belum.

*“White-box testing* yaitu metode desain *test case* yang menggunakan stuktur kontrol desain posedural unntuk memperoleh *test case”*[16]

*Test case* dapat diperoleh dengan:

1. Menjamin bahwa semua *independet path* pada suatu modul telah digunakan minimal satu kali
2. Menggunakan semua keputusan logis pada sisi *true* dan *false.*
3. Mengeksekusi semua *loop* dalam batasannya dan pada batas operasionalnya
4. Mengunakan struktur data internal untuk menjamin validitasnya



**Gambar 2.9** *White box Testing*

Sumber : https://www.slideserve.com/charo/teknik-pengujian-perangkat-lunak [17]

Pengujian ini harus memenuhi kriteria sebagai berikut :

1. Mengurangi pelaksanaan *test case* untuk mencapai hasil pengujian yang diinginkan
2. *Test case* akan menunjukkan ketidaksesuaian (ketidaksinkronan) beberapa kesalahan, tetapi kurang menunjukkan detil kesalahan

*Software* yang dibangun tidak selalu sempurna. Terkadang ada kesalahan yang baru akan diketahui saat pengujian atau bahkan saat implementasi. Berikut ini kriteria *software* yang “cacat” :

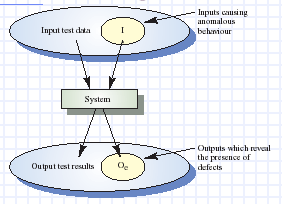
1. Kesalahan logika dan asumsi yang salah berbanding terbalik dengan probabilitas jalur program yang akan dieksekusi. Kecenderungan kesalahan ini terjadi pada tahap desain dan implementasi fungsi, kondisi atau kontrol yang berada di luar pikiran. Kesalahan ini tejadi jika pemrosesan yang rutin sudah dikerjakan dengan baik tetapi pemrosesan yang khusus cenderung diabaikan.
2. Selalu ada keyakinan bahwa *logical path* tidak akan dieksekusi pada basis regular. Kesalahan ini terjadi karena adanya kesalahan asumsi tentang aliran data dan kontrol.
3. Kesalahan tipologis yang merupakan kesalahan yang acak atau random. Perpindahan dari bahasa pemrograman satu ke bahasa pemrograman lain menyebabkan timbulnya kesalahan sintak.

*Path Testing* adalah metode pengujian yang memungkinkan desainer *test case* mengukur kompleksitas logika desain posedural dan menggunakannya sebagai pedoman untuk menetapkan basis set dari jalur eksekusi. Metode ini menjamin bahwa setiap statement akan dilalui minimal sekali dalam proses pengujian. Tujuannya adalah meyakinkan bahwa himpunan *test case* akan menguji setiap *path* pada suatu program paling sedikit satu kali. Titik awal *path testing* adalah suatu program *flow graph*  yang menunjukkan node – node yang menyatakan keputusan program, misalnya : kondisi if – then – else dan busur yang menyatakan alur kontrol. Statement yang memiliki kondisi ada pada node – node dalam *flow graph*. *Path testing* menggambarkan alur kontrol di mana setiap cabang ditunjukkan oleh path yang terpisah sedangkan loop ditunjukkan oleh arrows looping dan kembali ke loop kondisi node [16].

**2.2.12.2 *Black Box***

Iskandaria (2012), “Pengujian blackbox (blackbox testing) adalah salah satu metode pengujian perangkat lunak yang berfokus pada sisi fungsionalitas, khususnya pada input dan output aplikasi (apakah sudah sesuai dengan apa yang diharapkan atau belum). Tahap pengujian atau testing merupakan salah satu tahap yang harus ada dalam sebuah siklus pengembangan perangkat lunak (selain tahap perancangan atau desain)”[18]

Black-box testing berfokus pada persyaratan fungsional perangkat lunak. Metode ini memungkinkan *software developer* untuk mendapatkan serangkaian kondisi input yang mempergunakan semua persyaratan fungsional program. Black-Box testing bukan alternatif *white-box testing*, namun merupakan pelengkap yang mampu mengungkap kesalahan, jika dibandingkan metode *white-box testing.*



**Gambar 2.10**. *Black Box Testing*

**Sumber** : https://www.slideserve.com/charo/teknik-pengujian-perangkat-lunak[17]

Pengujian black box berusaha menemukan kesalahan dalam kategori :

1. Fungsi – fungsi yang tidak benar atau hilang
2. Kesalahan *interface*
3. Kesalahan dalam struktur data atau akses database eksternal
4. Kesalahan kinerja
5. Inisialisasi dan kesalahan terminasi

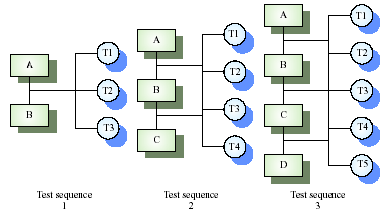
*Black-box testing* cenderung diaplikasikan selama tahap akhir pengujian. Pengujian black-box memperhatikan struktur kontrol, sehingga perhatian berfokus pada domain informasi. Black-box testing testing dirancang untuk dapat menjawab pertanyaan berikut :

1. Bagaimana validitas fungsional diuji?
2. Kelas input apa yang akan membuat test case menjadi baik?
3. Apakah sistem sensitif tehadap nilai input tertentu?
4. Bagaimana batasan dari suatu data diisolasi?
5. Bagaimana kecepatan dan volume data yang dapat ditolelir oleh sistem?
6. Apa pengaruh kombinasi tertentu dari data terhadap operasi sistem?

*Equivalence partitioning* yaitu metode pengujian *black-box* yang membagi domain input dari suatu program ke dalam kelas data dari mana *test case* dapat dilaksanakan. Input data dan output hasil terdapat di kleas yang berbeda yang sesuai dengan kelas inputnya. Masing-masing kelas *equivalensi partition* diproses dimana program akan memproses anggota kelas-kelas tersebut secara equivalen. *Test cases* dipilih dari masing-masing partisi

Partisi Ekivalensi (*equivalensi partition*) adalah Input data dan output hasil terdapat di kelas yang berbeda yang sesuai dengan kelas inputnya. Masing-masing kelas *equivalensi partition*  dimana program akan memproses anggota kelas-kelas tersebut secara *equivale*

*Integration testing* merupakan pengujian sistem secara keseluruhan atau subsistem yang terdiri dari komponen telah terintegrasi. Tes integrasi menggunakan black box dengan test caseditentukan dai spesifikasi. Kesulitannya adalah menemukan atau melokasikannya. Penggunaan *incremental integration testing* dapat mengurangi masalah tersebut.



**Gambar 2.11.** *Incremental Integration Testing*

**Sumber** : Https://www.slideserve.com/charo/teknik-pengujian-perangkat-lunak[17]

Ada dua pendekatan *integration testing* yaitu:

1. *Top – Down Testing,* berawal dari level atas sistem dan terintegrasi dengan mengganti masing – masing komponen secara *top- down*  dengan suatu *stub* (pogram pendek yang meng-*generate* input ke sub sistem yang diuji).
2. *Bottom – up Testing,* mengintegrasi komponen di level hingga sistem lengkap sudah teruji. Pada prakteknya, kebanyakan test integrasi menggunakan kombinasi kedua strategi pengujian tersebut.

*Interface testing* dilakukan jika modul – modul dan sub sistem terintegrasi dan membentuk sistem yang lebih besar. Tujuannya adalah mendeteksi *fault* terhadap kesalahan interface atau asumsi yang tidak valid tentang inteface tersebut. Pengujian ini sangat penting untuk pengujian pengembangan berorientasi objek yang didefinisikan oleh objek – objeknya.

Ada empat tipe interface, yaitu :

1. *Parameter interface,* di mana data dikirim dari satu prosedur ke prosedur lainnya
2. *Shared memory interface,* di mana blok memori di-*share* di antara prosedur – prosedur.
3. *Procedural interface,* di mana sub sistem mengenkapsulasi sekumpulan prosedur – prosedur yang akan dipanggil oleh sub sistem lainnya.
4. *Message passing interface,* di mana sub sistem meminta service dai sub sistem lainnya.

Terdapat beberapa *interface erors,* yaitu :

1. *Interface Misuse,* yaitu jika komponen pemanggil memanggil komponen lainnya dan membuat suatu kesalahan dalam penggunaan interfacenya. Misalnya : parameter dengan urutan yang tidak sesuai.
2. *Interface missunderstanding,* yaitu jika komponen pemanggil salah dalam mengasumsikan komponen *behaviour* yang dipanggil.
3. *Timing errors,* yaitu jika komponen yang memanggil dan yang dipanggil beroperasi pada kecepatan yang berbeda sehingga dimungkinkan mengakses informasi yang tidak *up to date.* Kesalahan terjadi karena *synchronization problem.*

Langkah pertama jika akan melakukan testing interface adalah merancang tes di mana parameter ke prosedur yang dipanggil berada pada nilai batas ekstrim. Perlu diingat bahwa tes menggunakan pointer null. Kemudian perancangan proses dilakukan sehingga komponen yang dites akan *fail.* Bisa juga digunakan *stress testing* pada *message passing. Stress testing* adalah menguji sistem dengan nilai yang melebihi kapasitas maksimum sistem. Stressing suatu sistem dapat mengurangi kemudahan rusaknya sistem. Sistem seharusnya tidak gagal total. Stess melacak kehilangan *service* yang tidak diduga maupun data yang hilang. Khusus untuk sistem terdistribusi dapat menyebabkan degradasi jaringan sehingga *overload.* Kemudian langkah terakhir adalah memvariasikan urutan di mana komponen diaktifkan pada *shared memory system*.

* 1. **Kerangka Pemikiran**

**Masaalah**

1. Bagaimana cara mengetahui cluster komoditi unggulan daerah Provinsi Gorontalo menggunakan Algoritma K-Means.?
2. Bagaimana hasil penerapan Algoritma K-Means untuk mengetahui cluster dari komoditi unggulan daerah Provinsi Gorontalo.?

**Pemodelan**

Observasi dan Dokumentasi

Pengumpulan Data Set

Metode

*K Means*

**System Development**

Diagram Konteks, Diagram Berjenjang,DAD Level 0 dst. Kamus Data

Konstruksi Sistem

Desain Sistem

Analisis Sistem

Desain Output, Input, Basis Data, Teknologi, dan Program

Database*(MYSQl),*Program *(PHP)*

Program (*White Box Testing)*

*Interface (BlackBox Testing)*

Pengujian Sistem

**Tujuan**

1. Untuk mengetahui cara merekayasa aplikasi clustering komoditi unggulan daerah Provinsi Gorontalo menggunakan Algoritma K-Means.
2. Untuk Menerapkan Algoritma K-Means dalam Clustering komoditi unggulan daerah Provinsi Gorontalo

**Gambar 2.12** Bagan Kerangka Berpikir

**Gambar 2.12** Bagan Kerangka Berfikir