

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kemiskinan menjadi salah satu masalah yang dialami oleh sebagian banyak para Negara berkembang. Banyak cara untuk menangani permasalahan kemiskinan salah satunya diantaranya antara lain mengadakan bantuan sosial kepada seseorang yang tidak mampu.

dalam tingkatan kehidupan kelompok terdapat kondisi kehidupan yang tidak sesuai dengan tingkatan kehidupan tertentu yaitu seseorang tidak mampu merawat kehidupanya sendiri. Serta tidak efektif dalam memakai tenaga mental, maupun fisiknya dalam kelompok tersebut yang dapat diartikan sebagai kemiskinan (Soekanto, 1982) yang diacu oleh (Setyawan, 2018). Menyadari hal tersebut, pemerintah membuat kebijakan baru yang berkaitan dengan rakyat miskin. dari kebijakan tersebut berupa Program Keluarga Harapan (PKH).

Program Keluarga Harapan (PKH) salah satu program bantuan dana tunai bersyarat pertama. Program ini diluncurkan pada Tahun 2007 di Negara Indonesia. Sesuai pengertian diatas Program PKH bertujuan untuk memberikan sebuah bantuan dana yang bersifat bersyarat. atau dalam artian lain untuk mendapatkan bantuan dari pemerintah, calon Penerima PKH harus menyelesaikan pesyaratan dari pemerintah yang berupa mengakses layanan Kesehatan dan Pendidikan tertentu.

Dalam menentukan calon penerima PKH terdapat kerumitan dalam pengolahan data selama ini yaitu dalam menentukan penduduk miskin yang menjadi prioritas pertama. Namun di desa semanding khususnya, data penduduk miskin sangat banyak. Oleh karena menentukan penduduk miskin sangat sulit dikarenakan data penduduk miskin sangat banyak. Selain itu, dalam proses penyaluran bantuan menjadi keluhan bagi masyarakat dimana dalam penyaluran bantuan masih dianggap belum tepat sasaran. Seperti halnya warga dapat dikatakan mampu mendapatkan bantuan sedangkan warga tidak mampu tidak mendapatkan bantuan. Hal itu membuat kecemburuan sosial antar warga di lingkungan masyarakat.

Dari Permasalahan tersebut, diperlukan adanya sebuah **Sistem Pendukung Keputusan penentu kelayakan penerima bantuan Program Keluarga Harapan (PKH) dengan menggunakan metode C4.5 dan membandingkan dengan metode K - Means**. Dengan tujuan menggunakan metode C4.5 dan K – Means yaitu terdapat Penelitian sebelumnya yang subjek yang sama dan metodenya kebanyakan hampir sama oleh karena itu, saya memakai metode C4.5 dan K – Means dimana metode tersebut belum banyak digunakan. Selain itu nantinya Aplikasi ini digunakan untuk membantu mempermudah atau setidaknya dapat mengurangi kendala dalam menentukan kelayakan Penerima PKH.

1.2. Perumusan Masalah

Menurut penjelasan dari latar belakang diatas, maka dapat menulis rumusan masalah yang akan diteliti dan akan dibahas di skripsi ini yaitu :

1. Bagaimana penerapan metode C4.5 dan K – means untuk penentuan kelayakan penerima Program Keluarga Harapan (PKH) ?
2. Bagaimana menentukan kelayakan penerima Program Keluarga Harapan dengan kriteria prioritas utamanya penduduk miskin dengan Algoritma *C4.5* dan *K - Means* ? dalam hal ini studi kasus data kebanyakan datanya terdapat penduduk miskin
3. Bagaimana analisis data untuk menentukan kelayakan calon penerima Penerima Program PKH menggunakan Algoritma *C4.5* dan *K – Means* ?

1.3. Batasan Masalah

Untuk menghindari permasalahan yang terlalu melebaran dari pokok masalah dan sistem pendukung keputusan penentu kelayakan penerima PKH menggunakan metode C4.5 dan K – Means berbasis web, agar tetap fokus di sistem tersebut. Maka Permasalahan dalam penelitian ini di batasi sebagai berikut :

1. Data yang akan digunakan menggunakan data yang diambil dari Seseorang Pendata PK Kelurga 2021 dan SGDS yang menerima PKH tahun 2018 di desa Semanding
2. Kriteria yang akan digunakan dalam metode C4.5 ialah sumber air minum, ukuran rumah , lansia , Ibu Hamil, Anak Sekolah, Jenis Lantai,

Luas Lantai, jenis dinding, jenis Atap. Dari sekian kriteria tersebut, kriteria tersebut telah dilakukan pembobotan dengan menggunakan rumus dalam menentukan kelayakan Penerima PKH.

3. Dalam menentukan penerima Program Keluarga Harapan (PKH), Aplikasi akan menghasilkan berupa layak dan tidak layak dalam menentukan penerima menggunakan metode C4.5 dan K – Means.
4. Bahasa pemrograman menggunakan *framework* php Laravel dan Database MySQL

1.4. Tujuan Penelitian

Menurut perumusan masalah dan Batasan Masalah telah dijelaskan sebelumnya, maka tujuan dari penelitian yang akan dilaksanakan dengan sebagai berikut :

1. Mengimplementasikan Algoritma C4.5 dan K – Means ke dalam Sistem Pendukung Keputusan menentukan kelayakan penerima Program Keluarga Harapan (PKH).
2. Menganalisis data penentuan kelayakan penerima Program Keluarga Harapan (PKH) menggunakan Algoritma C4.5 dan K – Means .
3. Membangun sistem informasi yang mengolah data dengan penerapan algoritma dari data mining yakni Algoritma C4.5 dan K – Means sebagai perkembangan kombinasi teknologi dengan ilmu data mining

1.5. Manfaat Penelitian

Dengan adanya penilitian ini mengharapkan dapat memberikan kesan positif untuk mengembangkan model Sistem Pendukung Keputusan Penentu Kelayakan Penerima Program Keluarga Harapan (PKH) Menggunakan Metode C4.5 dan K – Means. Dengan menggunakan model ini dapat memberikan cara yang lain dalam menentukan kelayakan calon penerima Program Keluarga Harapan (PKH) terhadap seseorang kebutuhan hidupnya kurang mampu.

Selain itu untuk mahasiswa, mengaplikasikan ilmu pengetahuan yang telah dipelajari untuk menolong atau membantu mengatasi ke dalam suatu masalah yang ada. Dan juga sebagai syarat kelulusan dalam menempuh Pendidikan Strata 1.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1. Tinjauan Pustaka

Pada Penelitian ini terdapat beberapa jurnal dan skripsi penelitian terdahulu sebagai acuan dan sebagai bahan pertimbangan yang berkaitan dengan penelitian yang akan dilteliti saat ini. Hal ini dilakukan karena untuk mengkaji penelitian yang akan dilakukan . dan juga untuk menambah wawasan pengetahuan dan teori. Dalam bagian ini penulis akan menampilkan beberapa jurnal dari penelitian sebelumnya dalam bentuk tabel.

Tabel 2.1 Penelitian Terkait

Jurnal I	
Judul	Implementasi Data Mining Untuk Manajemen Bantuan Sosial Menggunakan Algoritma K - Means
Penulis	Ali Ikhwan , Nuri Salim (2020)
Masalah	Terdapat Kesulitan dalam menentukan penduduk miskin yang datanya sebagian besar banyak data penduduk miskin.
Hasil	Dari permasalahan diatas, Teknik Clustering K – Means dapat digunakan untuk mengkaji permasalahan yang ada yakni memilih kelompok prioritas bagi orang berpenghasilan rendah untuk menerima bantuan.
Jurnal II	
Judul	Perbandingan Algoritma <i>C4.5</i> dan <i>Naïve Bayes</i> Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan
Penulis	Eka Fitriani (2020)
Masalah	Permasalahan yang terdapat pada bantuan PKH yakni seringnya adanya tidak tepatnya sasaran warga yang menerima Bantuan PKH.

Jurnal II (lanjutan)	
Hasil	Setelah dilakukan pengujian diantara metode yang berbeda di dapatkan hasil yaitu antara algoritma C4.5 dan Naïve Bayes, nilai keakurasan paling tinggi di dapat algoritma C4.5. dan di dalam jurnal tersebut menyatakan bahwa algoritma C4.5 salah satu metode yang cukup baik.
Skripsi I	
Judul	Sistem Informasi Geografis Pemetaan Warga Miskin Yang Pantas Mendapatkan Bantuan Menggunakan Metode Clustering K - Means
Penulis	Ugik Setyawan (2018)
Masalah	Menurut latar belakang dari skripsi tersebut, permasalahan yang terjadi adanya kurangnya informasi mengenai bantuan terhadap warga miskin. Sehingga penanganan warga miskin kurang begitu maksimal.
Hasil	Didapatkan sistem yang dapat mengolah data masyarakat miskin dengan mengelompokkan berdasarkan kriteria yang ada untuk memproses perhitungan. Juga menghasilkan pemetaan menggunakan Google Maps API yang di aplikasikan ke dalam sistm
Skripsi II	
Judul	Penerapan Metode C4.5 Untuk Penentuan Kelayakan Penerima Program Keluarga Harapan
Penulis	Irfan Abadi Saragih (2020)
Masalah	Penentuan kelayakan penerima Program Keluarga Harapan (PKH) masih dilakukan secara manual. Sehingga mempengaruhi hasil penentuan penerima pkh menjadi kurang maksimal.
Hasil	Dari hasil penelitian tersebut, telah dibuatkan suatu sistem yang dapat membantu Dinas Sosial yang terkait untuk mempermudah penentuan kelayakan penerima Program Keluarga Harapan (PKH).

2.2. Dasar Teori

2.2.1. Sistem Pendukung Keputusan

Sistem menurut (Kusrini,2007) yang diacu oleh (Tahyudin, 2014) sistem suatu group elemen dimana yang saling berkaitan dan bertanggung jawab dalam meproses suatu *input* sehingga menghasilkan berupa *output*.

Sistem dapat diartikan suatu jaringan kerja dari tata cara (procedure) yang saling berhubungan dan bersama sama melakukan tugasnya atau melakukan suatu kegiatan tertentu dan dapat menyelesaiakannya. Pendekatan lebih pada sistem yang mendefinisikan sistem sebagai kumpulan elemen – elemen atau component untuk mencapai suatu tujuan (Jogiyanto,2005) yang diacu oleh (Tahyudin, 2014).

Sistem Pendukung Keputusan dapat diartikan sistem informasi interaktif yang mana terdapat informasi berupa kumpulan data yang akan diproses atau diolah yang menghadirkan data baru yang nantinya dapat menentukan keputusan berdasarkan pemrosesan data baru tersebut. Selain itu, Sistem Pendukung Keputusan dapat mengambil keputusan semi terstruktur dan situasi tidak terstruktur. (Alter,2002) yang diacu oleh (Tahyudin, 2014).

Aplikasi DSS dapat digambarkan sebagai suatu sistem dibangun untuk yang memberikan dukungan yang berupa solusi dalam suatu masalah untuk mengevaluasi suatu peluang dalam permasalahan yang ada. Aplikasi DSS menggunakan CBIS (*Computer Based Information System*) dengan flexibility, interaktif dan mudah diadaptasi, yang dapat mendukung solusi atas permasalahan yang tidak terstruktur. DSS lebih di arahkan untuk memberi sebuah dukungan dalam melakukan pekerjaanya berdasarkan analitis dalam keadaan atau situasi yang terstruktur dengan kriteria yang belum jelas. DSS juga memberikan userinterface interaktif yang memungkinkan dalam melakukan pengambilan keputusan dengan cara berbagai analisis yang menggunakan model – model yang tersedia (Tahyudin, 2014, p. 3).

2.2.1.1. Tujuan dari Sistem Pendukung Keputusan (menurut Turban, 2005):

1. Kecepatan komputasi. Komputer memungkinkan pengambil keputusan melakukan banyak komputasi secara cepat dengan biaya yang rendah

2. Dukungan kualitas. Dengan spesifikasi computer memungkinkan bisa meningkatkan kualitas keputusan yang dibuat .
3. Mengatasi keterbatasan psikologis dalam pemrosesan dan penimpanan
4. Meningkatkan Produktifitas. Membangun suatu relasi antar kelompok pengambilan keputusan dengan para pakar , bisa membuat harganya sangat mahal.

2.2.1.2. Menyelesaikan suatu masalah dengan Keputusan, dapat dilihat dari kestrukturannya dalam menentukan keputusannya dapat dibagi menjadi (Tahyudin,2014) :

1. Keputusan terstruktur (*structured decision*)

Keputusan terstruktur dimana suatu keputusan yang dilakukan bersifat rutin dan secara berulang – ulang. Tata cara Pengambilan keputusan sangat jelas, terutama dilakukan pada manajemen tingkat bawah.

2. Keputusan semiterstruktur (*semistructured decision*)

Keputusan semiterstruktur yakni agak berbeda dengan Keputusan terstruktur. Keputusan terstruktur memiliki dua sifat. Dimana sifat yang pertama Sebagian keputusannya bisa dilakukan dengan oleh sistem komputer. Sedangkan sifat kedua yakni sisanya yang lain harus dilakukan oleh pengambil keputusan. dalam pengambilan keputusan sesuai prosedur secara garis besar sudah ada, akan tetapi beberapa hal yang harus memerlukan kebijakan dari pengambil keputusan. Semacam keputusan ini diambil oleh para manajer di dalam suatu Perusahaan atau organisasi yang levelnya menengah contoh keputusan jenis ini antara lain sebagai pengevaluasian kredit, serta penjadwalan produksi dan pengendalian persediaan.

3. Keputusan tak terstruktur (*unstructured dicision*)

Keputusan tak terstruktur suatu Pengambil Keputusan yang sangat sulit penanganannya , hal ini dikarenakan tidak berulang – berulang dan hampir tidak terjadi. Keputusan tersebut harus berpengalaman dan berbagai sumber yang bersifat eksternal atau dari luar. Keputusan tersebut terjadi pada manajemen tingkat atas. Contohnya keputusan pengajuan pengembangan teknologi baru, dan keputusan perekurutan pegawai ke dalam perusahaan tertentu.

2.2.2. Data Mining

(Suyanto, 2019, p. 1) mengungkapkan bahwa “Data Mining” ilmu analisis yang memproses data menjadi penemuan data baru atau pengetahuan di dalam database (*knowledge discovery in database*) salah satu kependekan dari KDD. Pengetahuan tersebut dapat berupa hubungan kevalidan data kesatu dengan yang lainnya .

Data Mining ditujukan diambil bagian intisarinya dari data tersebut dari sebuah group (Perkumpulan) data, akhirnya mendapatkan sebuah struktur data yang dipahami oleh manusia. Yang meliputi pra – pemrosesan data, pertimbangan model, ketertarikan ukuran, kompeksitas, setelah pemrosesan terhadap struktur telah ditemukan (pengetahuan), Penampilan dan update data secara *online* (Suyanto, 2019, p. 2).

2.2.2.1. Metode Data Mining

Data mining memiliki beberapa metode yang dapat memecahkan persoalan dalam proses pencarian data baru atau pengetahuan didalam database dan bisa digunakan setiap fungsinya. Berikut beberapa teknik yang dapat digunakan untuk memecahkan permasalahan di *data mining* antara lain (Nofriansyah & Nurcahyo, 2019):

1. Estimasi

Estimasi ini digunakan untuk melakukan sesuatu perkiraan terhadap data baru yang tidak memiliki keputusan yang berdasarkan dari data yang ada. Sebagai contohnya estimasi pembiayaan RAB terhadap pembangunan sebuah Gedung Rumah Sakit dengan acuan harga RAB setiap harganya beda kota beda harga.

2. Asosiasi

Asosiasi ini digunakan untuk mengetahui perilaku terhadap kejadian kejadian khusus sebelumnya. di dalam proses yang mana asosiasi muncul setiap kejadian. Terdapat metode untuk memecahkan permasalahan ini dengan menggunakan Algoritma Apriori yang sering digunakan. Sebagai Contoh pemanfaatan metode Apriori yakni menempatkan tata letak sebuah penjualan produk berdasarkan produk yang sering di beli oleh konsumen yang berada di

minimarket. Selain itu, menempatkan tata letak sebuah buku di rak buku yang berada perputakaan.

3. *Klasifikasi*

Suatu metode yang dilakukan dengan melihat perilaku pada atribut dari sebuah kelompok data yang telah didefinisikan. Cara ini memberikan klasifikasi terhadap data baru dengan cara mengedit (memanipulasikan) sebuah data yang ada telah diklasifikasikan . Algoritma Klasifikasi yang paling populer saat ini yaitu *Decision Tree* menjadi metode pengklasifikasian yang paling populer yang dapat di Implementasikan seperti *Algoritma C4.5*, ID3 dan lain sebagainya. Contoh Metode pengklasifikasian yang dapat kita temui yakni pengkelasan terhadap siswa yang mampu masuk kedalam kelas unggulan.

4. *Klustering*

Suatu metode dengan cara pengkelompokan terhadap sebuah dataset. Atau dapat diartikan menganalisis pengklompokan yang berbeda terhadap sebuah data. hampir mendekati dengan klasifikasi, tetapi metode klustering (pengklompokan) belum terdefinisikan berbeda dengan menggunakan klasifikasian. Algoritma yang cocok dengan metode klustering yakni antara lain : *Algoritma K – Means*, *Neural Network* , *Analitcs hierarki cluster*. Clustering membagi item menjadi kelompok – kelompok yang berdasarkan yang ditemukan tool data mining.

5. *Prediksi (Forecasting)*

Prediksi suatu metode yang dapat memperkirakan suatu peristiwa yang akan terjadi namun belum tentu pasti terjadi. Contohnya Badan Meterologi Klimatologi dan Geofisika (BMKG). Bagaimana BMKG dapat meramalkan Cuaca, Apakah Panas , Hujan, dan lain sebagainya pada daerah atau kota tertentu. Ada beberapa metode yang dapat digunakan salah satunya yaitu Metode Rough Set.

2.2.2.2. Tahapan Proses Data Mining

Dari dalam jurnal “*From Data Mining to Knowledge Discovery in Database*”, (Fayyad et al., 1996). Mengatakan bahwa data mining

salah satu tahapan yang ada di dalam seluruh proses KDD yang mana terdiri dari aplikasi analisis data dan *Algoritma* (metode) pencarian yang menghasilkan sebuah

pengetahuan (pola tertentu) dari sebuah data. Kedua istilah tersebut kerap digunakan tetapi memiliki konsep yang berbeda dan berkaitan dengan satu sama lain. Pada *Proses Knowledge Discovery Database (KDD)* terdapat beberapa tahapan yakni sebagai berikut (Saragih, 2020).

1. *Selection*

Selection data yaitu memilih data tertentu dengan kriteria atribut tertentu dari sekumpulan data yang harus dilakukan sebelum tahap proses menggali informasi dalam *Knowledge Discovery in Database* kependekan dari KDD dimulai. setelah proses penyeleksian terhadap sekumpulan data maka disimpan ke dalam suatu berkas, file terpisah .data tersebut nantinya yang akan digunakan untuk proses data mining.

2. *Preprocessing Data*

Suat tahapan di mana *Preprocessing data* mencakup beberapa hal yakni menghapus data ganda, memeriksa data yang salah (inkonsisten), dan memperbaiki kesalahan data, contohnya kesalahan ketik (tipografi). Serta melakukan proses enrichment, yaitu suatu tahapan proses “memperkaya” data yang sudah ada dengan informasi lain yang relevan dan diperlukan untuk KDD.

3. *Transformasi Data*

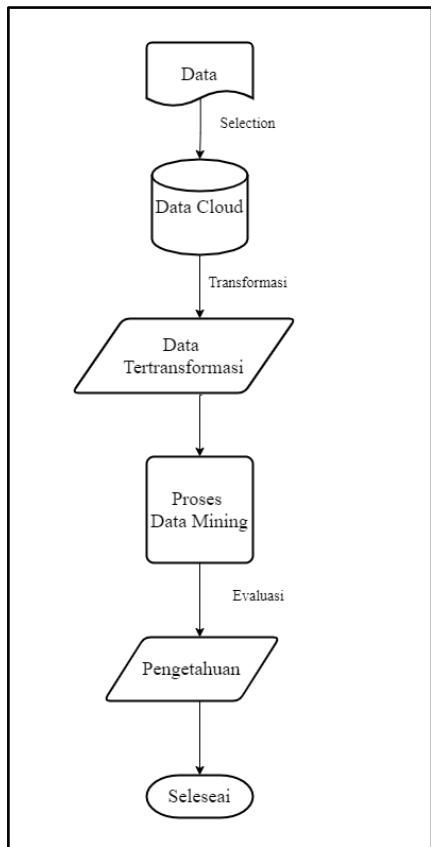
Pada fase ini yang dilakukan berupa Transformasi data dimana proses merubah data yang ada belum memiliki entitas (objek) yang jelas ke dalam bentuk data valid atau siap untuk dilakukan proses data mining.

4. *Data Mining*

Data Mining menjadi proses pencarian pengetahuan atau informasi menarik yang ada pada data terpilih dengan menggunakan metode tertentu.

5. *Interpretation (Evaluation)*

Pada tahap yang terakhir ini dilakukan proses yang perlu ditampilkan ke bentuk yang mudah untuk dipahami oleh pihak yang berkepentingan atau orang. Proses *Knowledge Discovery in Database (KDD)* dapat dilihat pada gambar 2.1 sebagai berikut



Gambar 2.1 Flowchart proses KDD

2.2.3. Algoritma C4.5

Salah satu seseorang Bernama J. Ross Quinlan menjadi peneliti bidang *machine learning*, pada akhir tahun 1970 sampai tahun 1980 awal. Dengan membuat sebuah Algoritma *dec* yang dapat artikan dalam istilah lain dari ID3 (*Iterative Dichotomiser*). Kemudian Ross mengembangkan Algoritma ID3 menjadi Algoritma C4.5 (*Decision Tree*), yang memiliki kelebihan dari Algoritma sebelumnya yakni fleksibel, dapat divisualisasikan bentuk gambar (pohon keputusan) juga dapat dimengerti. (Nasrullah, 2018, p. 245) mengatakan Algoritma C4.5 dapat artikan sebagai struktur dari pohon keputusan yang mana mempunyai simpul node yang menggambarkan atribut. Setiap cabang mendeskripsikan hasil pengujian dari atributnya, dan setiap daun menggambarkan kelasnya.

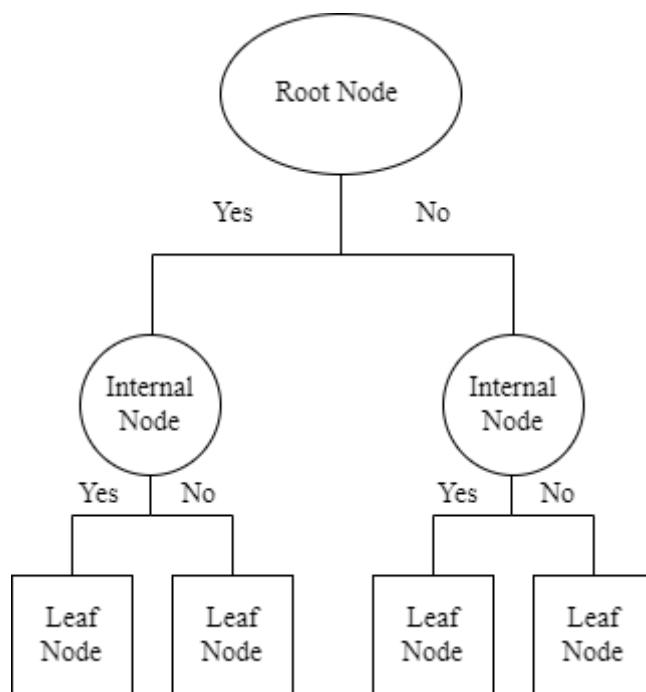
Salah satu metode klasifikasi yang menggunakan struktur pohon (tree) yang dapat dikatakan sebagai metode Algoritma C4.5 dengan pohon keputusan. Yang mana memiliki node sebagai atribut, cabangnya menggambarkan sebagai nilai dari

atribut, dan daunnya sebagai hasil pengujian dari atributnya (kelas). *root* salah satu node akar atau node yang paling atas dari pohon keputusan (Ardiansyah Sembiring et al., 2018, p. 75).

Menurut (Rufiyanto & dkk, 2021, p. 6) *Decision Tree* mengatakan bahwa algoritma klasifikasi data mining dengan membentuk pola pohon keputusan yang digunakan dapat memecahkan masalah yang diinputkan. Dengan menggunakan pola pohon keputusan, dapat menemukan suatu masalah antar faktor – faktor yang dapat memicu masalah dan melakukan pencarian solusi dengan memperhitungkan faktor – faktor tersebut.

Pohon keputusan mempunyai 3 jenis node, yaitu sebagai berikut (Saragih, 2020:13):

1. *Node Root*, dapat dikatakan sebagai node akar (paling atas) yang terdiri dari beberapa *output* atau bisa tidak mempunyai *output* dan tidak mempunyai *input*.
2. *Node Internal*, dapat diartikan node cabang, pada node ini terdapat satu *input* dan mempunyai *output* yang minimal dua.
3. *Terminal node*, node daun yang menggambarkan kelasnya, biasanya node ini terdapat satu *input* dan tidak mempunyai *output*.



Gambar 2.2 Node pada Dicision Tree

Decision tree menggunakan aturan *if – then*, juga tidak membutuhkan metrik dan parameter. *Decision tree* memiliki struktur sangat sederhana dan mudah di jelaskan. *Decision tree* juga dapat memecahkan masalah atribut *multi – type*, dan nilai nilai yang hilang dan data *noise* (Saragih, 2020).

Algoritma C4.5 menjadi algoritma yang menggunakan gain ratio untuk membentuk Decision Tree, oleh karena itu diperlukan beberapa tahapan perhitungan dalam menentukan Decision Tree sebagai berikut.(Gorunescu, 2011) di dalam bukunya (Rufiyanto & dkk, 2021):

1. Menyediakan *data training* atau *data sets* yang dapat diambil dari data *history* yang terjadi pada sebelumnya dan dikelompokkan ke dalam kelas – kelas tertentu.
2. Hitung terlebih dahulu nilai *index entropy*, nantinya untuk menentukan *root node* dari pohon keputusan dengan menghitung nilai *gain* tertinggi dari masing - masing atrribut atau opsi kedua menggunakan nilai *entropy* terendah. Berikut rumus *entropy* :

$$Entropy(S) = \sum_{i=1}^n -p_i \cdot \log_2 p_i \quad (2.1)$$

Keterangan :

S : himpunan kasus

n : jumlah partisi

Pi : proporsi Si terhadap S

3. Menghitung *Gain* untuk menentukan node akar dengan rumus sebagai berikut :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.2)$$

Keterangan :

S : himpunan kasus

A : Atribut

N : jumlah partisi atribut A

|Si| : Proporsi Si terhadap S

|S| : jumlah kasus dalam S

4. Untuk menghitung *gain ratio*, perlu diketahui harus menghitung split Information terlebih dahulu yakni dengan rumus sebagai berikut

$$Split\ Information = - \sum_{t=1}^c \frac{s_t}{s} \log_2 \frac{s_t}{s} \quad (2.4)$$

Keterangan :

S_1 sampai S_c = c subset salah satu hasil dari pemecahan S dengan Menggunakan atribut A mempunyai nilai sebanyak c

5. Selanjutnya menghitung gain ratio nantinya juga bisa digunakan menentukan node akar (jika data record bertipe numeric)

$$Gainratio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)} \quad (2.5)$$

6. Ulangi Langkah ke-2 hingga semua *record* (data) terpartisi

Proses pengisian *Decision Tree* akan berhenti jika :

1. Semua data dalam *record* mendapatkan kelas yang sama.
2. Tidak ada atribut dalam record di partisi lagi
3. Tidak ada cabang yang kosong di dalam *record*



Gambar 2.3 Algoritma Penyelesaian Algoritma C4.5

2.2.3.1. Jenis Jenis Algoritma Klasifikasi

Jenis Jenis Perkembangan Algoritma Decision Tree antara lain (Rufiyanto & dkk, 2021) :

- 1. CART (*Classification and Regression Tree*)**

Algoritma CART suatu Algoritma *decision tree* yang membentuk sebuah cabang dan juga membentuk node menggunakan perhitungan *GiniGain* yang berbasis *IndexGini*.

- 2. ID3 (*Iterative Dichotomiser 3*)**

Algoritma ID3 dapat dikatakan sebagai sebuah Algoritma *decision tree*, dalam proses perhitungannya menggunakan *Entropy* dan *information gain* yang akan digunakan untuk memilih *atribut* menjadi *node* akar yang berbasis *information gain*, pernyataan tersebut dari (Yan, Guo, Jin. 2018) didalam bukunya (Rufiyanto & dkk, 2021)

- 3. C4.5**

Algoritma C4.5 menjadi salah satu Algoritma *decision tree* pengembangan dari Algoritma ID3. Konsep sama seperti dengan ID3 yang menggunakan *Entropy* dan *Information Gain* untuk memilih *attribute* dijadikan *Node* Akar nantinya . dan memiliki kelebihan dari Algorima sebelumnya, yang awal fiturnya hanya bertipe kategorikal dan tipe *numerik* yang tidak bisa digunakan (Quinlan, 1996). Sedangkan C4.5 menangani fitur berupa pemotongan *decision tree* dan penurunan *rule set* (Rufiyanto & dkk, 2021).

2.2.4. Algoritma K – Means (Clustering)

Analisis Cluster salah satu bagian dari beberapa metode yang ada pada data mining yang digunakan untuk menemukan sebuah kelompok object, yang memiliki kemiripan karakteristik tertentu dan dipisahkan dengan kelompok lainnya, sehingga object lebih homogen (mirip) di dalam kelompok yang sama dari pada kelompok yang berbeda. Untuk jumlah kelompok yang dapat teridentifikasi tergantung banyaknya data dan berbagai macam Objek. Tujuan dari pengelompokan (*Clustering*) dari sekumpulan data set kedalam beberapa kelompok

yang masing – masing kelompok mempunyai sebuah karakteristik tertentu yang dapat dibedakan satu sama lainnya yakni untuk menganalisis dan penafsiran lebih lanjut dengan penelitian yang dilakukan (Nofriansyah & Nurcahyo, 2019, p. 52).

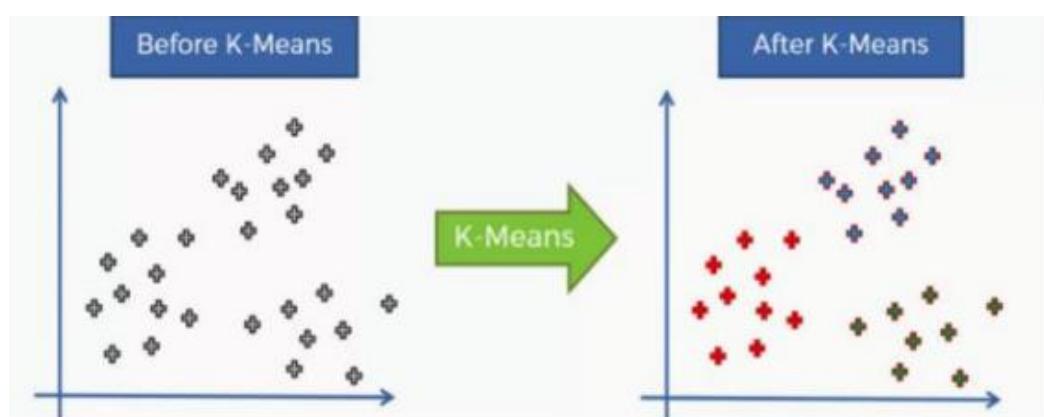
Metode *Clustering* akan membentuk sebuah kontruksi hirarki yang berdasarkan tingkatan (level) tertentu seperti struktur pohon yang menjadi pengertian dari *hierarchical methods* (Teknik Hiraki). Oleh karena itu dalam proses pengelompokannya dilakukan secara bertahap atau step by step (Nofriansyah & Nurcahyo, 2019, p. 55).

Algoritma K – Means ditemukan oleh beberapa orang yaitu McQueen pada tahun (1967), Forgey pada tahun (1965), Lloyd pada tahun (1957), Friedman dan Rubin pada tahun (1967). Metode Pengeklompokan (*Clustering*) di publikasikan pada tahun (1982), yang awalnya pertama kali ditemukan oleh Lloyd (1957). Namun pada tahun 1965 sebelumnya Forgey lebih dulu mempublikasikan teknik yang sama yang dikenal sebagai Lloyd-Forgy Primartha,2018) diacu oleh (Wanto et al., 2020).

K – Means salah satu Algoritma Pengelompokan (*Clustering*) yang termasuk dalam kelompok Unsupervised learning yang berfungsi untuk membagi sebuah dataset dengan beberapa kelompok dengan penyimpanan berbeda. Sedangkan dengan menggunakan Metode K – Nearest Neighbor (KKN) salah satu dari algoritma clustering termasuk kategori supervised learning yang menerima inputan berupa vektor. Metode K – means, pada komputer mengelompokkan inputannya berupa data dalam proses pengelompokan secara mandiri tanpa mengetahui tujuan kelasnya masing – masing. Dalam hal ini inputan yang diterima yaitu data atau objek dan k menjadi sebuah kelompok (Cluster) yang telah ditentukan (target). Dalam Pemrosesan data menggunakan Algoritma K – Means Clustering, akan dimulai dengan step pertama pengelompokan Centroid pertama di pilih secara acak menjadi titik awal pada setiap kelompok (Cluster). Lalu kemudian menghitung secara berulang ulang agar nantinya posisi Centroid Agar optimal(Wanto et al., 2020).

Secara sederhana Algoritma K – Means mengambil sebagian dari banyaknya data set untuk menjadikan sebuah pusat cluster awal, untuk menentukan

pusat (centroid) cluster ini dipilih secara acak dari data set. Selanjutnya algoritma K – Means akan mengetest masing – masing dari setiap bagian bagian (data) dalam data set, dan menandai data tersebut ke salah satu pusat (centroid) cluster yang sudah ditentukan (terdefiniskan) sebelumnya yang tergantung jarak terdekat antar masing – masing data dengan setiap pusat cluster . Selanjutnya posisi Cluster akan dihitung ulang sampai semua data digolongkan ke dalam tiap – tiap cluster dan terakhir akan membentuk cluster baru (Sihombing,2017) diacu oleh (Wanto et al., 2020).



Gambar 2.4 Konsep Clustering K – Means (Anjar wanto dkk,2020:4)

2.2.4.1. Penerapan Algoritma K - Means

Pengaplikasian Algoritma K – Means dapat di terapkan berbagai bidang antara lain (Setyawan, 2018, pp. 17–19):

A. Berdasarkan untuk Pemahaman (*Understanding*)

Clustering berdasarkan Pemahaman (*Understanding*) bertujuan untuk menghasilkan kelompok – kelompok yang meliputi objek – objek yang mempunyai karakteristik yang sama. Seperti halnya mengelompokkan objek – objek.

1. Penerapan dalam Bidang Biologi

Algoritma K – Means dapat di aplikasikan ke dalam bidang Biologi yakni bertujuan untuk mengelompokkan gen berdasarkan polanya. Oleh Karena itu, dibutuhkan untuk menemukan gen yang memiliki karakteristik (fungsi) serupa.

2. Penerapan dalam Bidang Bisnis

Algoritma K – Means juga diterapkan pada bidang bisnis yakni untuk melakukan segmentasi pasar. Segmentasi pasar dapat diartikan sebagai pengelompokan (*clustering*) customernya berdasarkan karakteristik mereka, misalnya : gaya hidup, kebutuhan (Setyawan, 2018, p. 18). Selain itu Algoritma K – means digunakan dalam sistem Pemberi keputusan (Rekomendasi) untuk mengelompokkan objek – objek yang terkait.

3. Penerapan dalam Bidang Temu Kembali informasi

Algoritma juga dapat digunakan dalam mengelompokkan dokumen berdasarkan topiknya keyword menggunakan sistem temu kembali informasi.

(Putung et al., 2016) Sistem Temu Kembali Informasi atau yang disebut *information retrieval system* salah satu sistem digunakan menemukan informasi yang bersangkutan yang dibutuhkan oleh penggunanya, dengan menerapkan sistem tersebut permasalahan akan terselesaikan dalam pencarian sebuah dokumen yang akan membeberikan hasil yang sesuai dibutuhkan oleh penggunanya.

B. Berdasarkan untuk *Utility*

Clustering berdasarkan *Utility*, digunakan untuk mengelompokkan sebuah himpunan data yang besar agar memudahkan analisis data untuk memproses data lebih lanjut. *Centroid*, dari *cluster* memegang peran lebih dalam hal ini

1. Penerapan di dalam Kompresi Data dan Multimedia

Algoritma K – Means digunakan kompresi data dan multimedia seperti halnya *citra, audio, video*. Setiap objek data (*pixel*) yang mempresentasikan dengan *centroid* dari sebuah *cluster* yang dimuat dari objek bersangkutan, yang diartikan sebagai pengertian dari kuantisasi vector.

2. Rangkuman Data

K – means juga digunakan dapat mengelompokkan data sebelum menerapkan metode analisis data yang ada seperti *regresi*, tetangga terdekat atau PCA. Dimana K – Means digunakan terlebih dahulu dalam mengelompokkan data ke dalam *cluster – cluster*. yang selanjutnya metode analisis data hanya diterapkan pada *centroid* dari setiap *Cluster* sehingga lebih efisien dalam penggunaan waktu dan ruang.

2.2.4.2. Karakteristik K - Means

Algoritma K – Means memiliki beberapa kelebihan dan kekurangan di dalam *Karakteristik K – Means*. berikut beberapa kelebihan dan kekurangan dari Algoritma k – means antara lain (Wanto et al., 2020, p. 6):

1. K – Means salah satu teknik pengelompokan yang secara sederhana juga mudah dan sangat cepat dalam proses penerapan *Clustering*
2. Pada sekumpulan sebuah data tertentu Algoritma K – Means dalam pengelompokan data tidak bisa dilakukan dengan baik, yang mana hasil pengelompokan tidak memberikan (output) pola kelompok yang mewakili karakteristik dari sekumpulan data tersebut.
3. Algoritma K – Means kemungkinan bisa terjadi masalah (error) Ketika mengelompokkan data yang mengandung data outlier. Dimana data Outliers salah satu sebuah data yang menyimpangnya terlalu jauh dari data yang lainnya dalam suatu rangkaian sebuah data.
4. Algoritma K – Means sangat peka (sensitif) dalam pembangkitan Centroid awal yang secara random.
5. Sebuah data tertentu Kemungkinan bisa terjadi suatu cluster tidak mempunyai anggota.
6. Setelah Proses Clustering dengan menggunakan Algoritma K – Means tidak bersifat unik selalu berubah – ubah kemungkinan bisa hasil baik, terkadang jelek.
7. Algoritma K – Means sangat sulit menentukan Nilai terbaik dari sebuah kumpulan data sets.

- Algoritma K – Means menyatakan bahwa Algoritma K – Means hanya dapat mengolah data kuantitatif. Kuantitatif dapat dikatakan sebuah data yang dapat dihitung secara langsung yang berupa informasi yang dinyatakan sebagai angka (Sugiyono,2010). Oleh karena itu Algoritma K – Means masuk kategori data kuantitatif.

2.2.4.3. Konsep K – Means

Cluseting mempunyai dua jenis data yang sering dipakai dalam proses pengelompokan yaitu *Hierarchical* dan *Non – Hierarchical*. Algoritma K – Means menjadi salah satu metode data Pengelompokan (*Clustering*) yang ada masuk ke dalam kategori *Non – Hierarchical (partitional Clustering)*. Konsep dari Algoritma K – Means yakni mengelompokkan sebuah data yang ada ke dalam beberapa kelompok. Di mana dalam dataset (satu kelompok) tersebut mempunyai beberapa karakteristik yang sama maupun berbeda antara satu dengan yang lain yang ada di dalam kelompok yang lain (Wantu et al., 2020, p. 5).

Berikut Langkah – Langkah menggunakan algoritma K – Means yang dapat di Jelaskan Sebagai Berikut (Wantu et al., 2020, pp. 5–6) :

- Menentukan jumlah cluster (k) pada sekumpulan data yang telah di normalisasikan
- Menentukan nilai pusat (Centroid) pada cluster, dimana Penentuan Centroid pada tahap awal dilakukan secara random (acak) sedangkan pada tahap iterasi menggunakan rumus seperti pada persamaan (2.6) berikut ini.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2.6)$$

Keterangan :

V_{ij} = Centroid rata – rata cluster ke – i untuk variabel ke – j

N_i = Jumlah anggota *cluster* ke – i

i, k = Indeks dari *Clustier*

j = Indeks dari variabel

X_{kj} = nilai data ke – k variabel ke – j untuk cluster tersebut.

3. Pada setiap data (*record*), hitung jarak terdekat dengan menggunakan *Centroid*. Untuk mengukur jarak sebuah data ke pusat kelompok, ada beberapa cara (metode) yang dapat digunakan antara lain : *Euclidean*, *Manhattan/City Block*, dan *Minkowsky*. Dari beberapa cara tersebut memungkinkan terdapat kelebihan dan kekurang masing – masing. namun untuk Penelitian Skripsi ini menggunakan *Euclidean Distance*, dengan rumus seperti dibawah ini:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (2.7)$$

Keterangan :

De = *Euclidean Distance*

i = Banyaknya objek – objek

(x,y) = Koordinat objek

(s,t) = Koordinat Centroid

4. Mengelompokkan objek (data) berdasarkan jarak ke Centroid terdekat
5. Mengulangi lagi step (Langkah) ke – 3 hingga ke – 4 , dan lakukan iterasi hingga *Centroid* Bernilai optimal

2.2.5. Program Keluarga Harapan

(Putri Pratiwi et al., 2019, p. 184) Program Keluarga Harapan (PKH) salah satu program yang ada dari pemerintah yang memberikan sebuah bantuan tunai Kepada Keluarga Penerima Manfaat (PKM), bantuan tersebut berbentuk bantuan bersyarat. Dimana Keluarga yang menerima program tersebut harus memenuhi syarat yang terkait dengan mengharapkan meningkatkaan kualitas Sumber Daya Manusia (SDM) yakni antara lain pendidikan dan kesehatan. Selain Pendidikan dan Kesehatan yang dijadikan syarat menerima Bantuan, PKH juga membuka akses bagi keluarga miskin terutama ibu hamil dan anak yang dimanfaatkan untuk berbagai fasilitas layanan kesehatan (faskes) dan fasilitas Pendidikan(fasdik) yang ada di sekitar lingkungan masyarakat tersebut. PKH sekarang juga di manfaat

dalam mempertahankan taraf kesejahteraan sosial seuai amanat konstitusi dan Wacita Presiden RI yang mencakup penyandang disabilitas dan lanjut usia.

Menurut (Adisanjaya Suleman & Resnawaty, 2017, p. 88) data Penduduk Miskin di Indonesia Miskin masih mengatakan cukup tinggi dan juga masih ada yang tergolong mendekati miskin. Tercatat jumlah penduduk miskin mencapai 28,01 juta orang atau dalam prosentase yakni 10,86 persen pada tahun 2016. Dengan adanya data tersebut, penduduk miskin masih sangat tinggi perlu ada gagasan mengadakan program perberdayaan yang efektif dan nantinya program tersebut dapat berjalan dengan baik dan teratur tanpa adanya diskriminasi oleh Pihak manapun. Namun nyatanya program pemerintah banyak terdapat kontroversi dan kritikan di kalangan masyarakat dalam bentuk yang diluncurkan ataupun penerapannya

Sebenarnya program dari pemerintah tersebut mempunyai tujuan yang baik, karena Program PKH tak jarang menyatakan bahwa program tersebut sebagai program pemberdayaan. oleh sebab itu program tersebut benar – benar memberikan akses (jalan) kepada anggota masyarakat yang tidak mampu (Putri Pratiwi et al., 2019, p. 89).

Menurut (Ife, 2002) didalam jurnalnya (Putri Pratiwi et al., 2019) menyatakan bahwa terdapat beberapa ciri – ciri keberdayaan yang dapat jelaskan sebagai berikut:

1. *power over personal choices and life chances*

yang dapat diartikan sebagai seseorang atau Sebagian orang kurang mampu dalam menentukan nasib mereka. contohnya seperti halnya orang yang memiliki ketidakmampuan pada saat mengambil suatu keputusan mengenai hidup mereka (Pada Kasus Kemiskinan).

2. *Power over the assertion of human right*

Dalam arti lain keberdayaan yang penyelenggaraan hak asasi manusia terhadap diri sendiri dan lingkungannya. Hal ini jika dihubungkan dengan program PKH, maka pemenuhan kebutuhan hidup yaitu salah satu hak asasi bagi

masyarakat yang ada di negara Indonesia. Namun Hambatanya dalam memenuhi kebutuhan, tidak terpenuhinya hak asasi manusia itu sendiri.

3. *Power over the definition of need*

Atau yang disebut juga Keberdayaan dalam mendefinisikan kebutuhan. Dimana dalam melalui perencanaan pembangunan pemerintah tidak ikut andil dalam menentukan kebutuhan masyarakat yang kurang mampu sehingga mengakibatkan ketidak sesuaiannya kebutuhan mereka.

4. *Power over ideas*

Dengan kata lain kerberdayaan dalam mengemukakan gagasan. Yang mana kemampuan seseorang dalam menyampaikan sebuah gagasan baik di forum, seseorang terlibat percakapan(dialog) dengan orang lain dan seseorang tersebut mempunyai ide yang berkontribusi dalam perbaikan hidup

5. *Power over institutions*

Ketidakmampuan masyarakat atau seseorang yang disebabkan oleh institusi sosial dalam hal antara lain sistem pendidikan, struktur pemerintahan. Hal ini menjadi Fokus dari program PKH.

6. *Power over resource*

Banyak sekali masyarakat yang tidak memiliki keberdayaan untuk menggunakan sumber sumber yang ada (kehidupan). Hal ini dapat terjadi baik untuk financial maupun nonfinancial resource seperti Pendidikan.

7. *Power over economic activity*

kemampuan seseorang atau komunitas untuk mengontrol mekanisme baik untuk proses produksi, maupun distribusi dan pertukaran yang terjadi. Hal tersebut berkaitan untuk melakukan usaha.

8. *Power over reproduction*

Suatu pemberdayaan yang tidak hanya proses produksi saja (kelahiran), namun juga dari mana sudut pandang seseorang sesuai latar belakangnya bagaimana ia memproses baik dalam hal sosialisasi dan pembelajaran yang dilakukan

2.2.6. Alat Bantu Analisis dan Perancangan Sistem

Dalam sebuah analisis dan perancangan Sistem Informasi, pasti ada membutuhkan beberapa perlengkapan (tools) yang nantinya digunakan untuk membantu perancangan nantinya agar analisa hasilnya tercapai. Dalam hal ini

perlengkapan yang dimaksud yaitu alat yang nantinya dapat membantu merancang setiap proses membangun sebuah sistem (Saragih, 2020).

2.2.6.1. UML

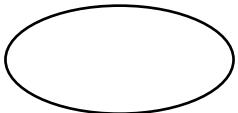
UML kepanjangan dari (*United Modelling Language*) salah satu Bahasa permodelan yang digunakan untuk suatu sistem atau perangkat lunak (software) yang menggunakan paradigm (pola) yang berorientasi objek. Modeling sebenarnya meringkas (penyerdahaan) permasalahan – permasalahan yang kompleks sedemikian sehingga dapat mudah di pelajari dan dipahami (Fatmawati & Munajat, 2018)

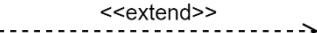
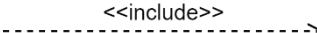
(Randa, 2018) menyatakan bahwa UML mempunyai beberapa diagram yang berbagai sudut pandang menurut kebutuhan dari perangkat lunak (sistem) yang terdiri dari sebagai berikut :

A. Use Case Diagram

Use case diagram mempresentasikan fungsionalitas yang dari sebuah sistem yang diharapkan nantinya. Dimana Use Case mendeskripsikan sebuah interaksi antara satu atau lebih actor dengan sistem Informasi yang dibuat nantinya. Berikut simbol Use Case diagram yang dapat dilihat pada Tabel 2.2

Tabel 2.2 Simbol Use Case Diagram (Umar Al Faruq, 2015 diacu oleh (Heriyanto, 2018, p. 16))

Simbol	Keterangan
Use case 	Use case sebuah fungsionalitas yang disediakan sistem sebagai komponen yang saling bertukar pesan atau informasi antar komponen atau actor. Yang biasanya menggunakan kata kerja di awal frase
Actor 	Seseorang atau orang yang berinteraksi dengan sistem yang dibuat, walaupun simbol dari actor yang mpresentasikan gambar orang , tetapi actor belum tentu

Simbol (Lanjutan)	Keterangan (Lanjutan)
	Orang . dan dinyatakan sebagai kata benda di awal frase nama actor.
Asosiasi 	digunakan untuk Mengkomunikasi antara <i>actor</i> dan <i>use case</i> yang saling terkait pada <i>use case</i> atau <i>use case</i> memiliki interaksi dengan <i>actor</i> .
Ekstensi 	Hubungan <i>use case</i> tambahan ke <i>use case</i> yang mana <i>use case</i> yang ditambahkan dapat berdiri sendiri walaupun tanpa <i>use case</i> tambahan.
Generalisasi 	Relasi generalisasi dan spesialisasi (khusus- umum) di mana antara dua buah <i>use case</i> yang salah satu fungsinya bersifat lebih umum dari lainnya.
Include 	Hubungan <i>use case</i> tambahan ke sebuah <i>use case</i> yang mana yang ditambahkan membutuhkan <i>use case</i> ini untuk menjalankan fungsinya atau sebagai syarat dijalankan <i>use case</i> ini.

B. Activity Diagram

Activity Diagram mendeskripsikan berbagai alir (Flow) aktivitas dalam sistem yang sedang dirancang, bagaimana awal sistemnya dan akhirnya, keputusan (decision) yang mungkin terjadi. Berikut simbol Activity Diagram yang dapat dilihat pada tabel 2.

Tabel 2.3 Simbol Activity Diagram (Urva & Fauzi Siregar, 2015)

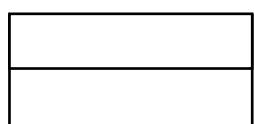
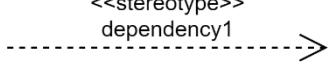
Simbol	Keterangan
Start Point	digunakan di awal aktivitas

Simbol (Lanjutan)	Keterangan (Lanjutan)
	Biasanya diletakkan pada pojok kiri atas
End Point 	Dinyatakan sebagai akhir dari aktivitas
Activity 	Dinyatakan sebagai proses atau pengolahan dan masing-masing kelas saling berinteraksi dengan yang lain
Fork 	Fork atau Percabangan digunakan untuk menunjukkan kegiatan (step) yang dilakukan secara parallel.
Join 	Join atau Penggabungan atau rake menyatakan untuk menggabungkan dua kegiatan parallel menjadi satu.
Decision 	<i>Decision</i> mempresentasikan untuk pengambilan sebuah keputusan dalam proses (kegiatan) nilai dari <i>decision</i> berupa boolean <i>true or false</i>
Swimlane 	Swimlane digunakan sebagai pembagian activity diagram untuk menunjukkan siapa? melakukan apa?.

C. Class Diagram

Class Diagram mempresentasikan struktur struktur sistem dari pendefenisian kelas yang dibuat untuk membuat sistem. Berikut Simbol class diagram yang dapat dilihat pada Tabel 2.4.

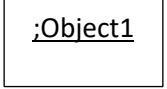
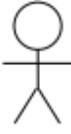
Tabel 2.4 Simbol Class Diagram (Rosa dkk, 2013)

Simbol	Keterangan
Generalization 	Hubungan antarkelas dimana memaksa generalisasi – spesialisasi (umum – khusus)
Nary Association 	Berusaha menghindar dari assosiasi dengan lebih dari dua objek
Class 	Kelas menjadi sebuah kumpulan suatu objek pada struktur sistem
Realization 	Suatu Proses (operasi) yang dilakukan oleh suatu objek.
Dependency 	Hubungan antar kelas dengan maksud saling ketergantungan dengan kelas lainnya
Association 	Hubungan antarkelas yang bermakna umum, assosiasi juga biasanya disertai dengan multiplicity

D. Sequence Diagram

Sequence Diagram ini menggambarkan sebuah interaksi antara objek yang didalam maupun disekitar sistem (termasuk pengguna, tampilan, dan lain sebagainya) yang dapat berupa message (pesan) yang digambarkan terhadap waktu. Berikut simbol class diagram yang dapat dilihat pada Tabel 2.5 :

Tabel 2.5 Simbol Sequence Diagram

Simbol	Keterangan
Lifeline 	Lifeline mengemukakan bahwa kebaradaan sebuah object dalam base waktu dimana Notasi lifeline mempunyai garis putus – putus vertical yang ditarik dari sebuah object
Activation 	Activation mempresentasikan sebuah kotak segi empat yang digambar pada sebuah lifeline, yang digunakan untuk mengindikasikan object yang akan melakukan sebuah aksi.
Message 	Message dinotasikan sebagai dengan anak panah horizontal diantara activation, message diindikasikan oleh komunikasi antara object – object
Object 	Object salah satu instance dari sebuah class dan ditulis secara horizontal. Dipresentasikan sebagai class (kotak) dengan nama object yang diawali dengan sebuah titik koma
Actor 	Actor dapat berinteraksi dengan suatu object

Sumber (Nurdam,2014, diacu oleh (Heriyanto, 2018)

2.2.6.2. Context Diagram

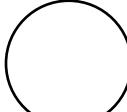
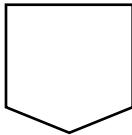
Context Diagram terdiri dari suatu proses yang mendeskripsikan ruang lingkup pada sistem dan menjadi salah satu langkah pertama dalam pembuatan DFD. Dalam context diagram menjelaskan sistem apa yang akan dibuat dan entity eksternal apa saja yang terliba. Juga context diagram harus ada arus data yang masuk maupun keluar (Wijaya et al., 2013).

2.2.6.3. Flowchart

Flowchart atau disebut sebagai bagan alur menjadi sebuah tahapan (proses) yang mempresentasikan alur (*flow*) dan bagan (*chart*) salah satu definisi dari flowchart. Yang mana digunakan untuk terutama dalam alat bantu komunikasi dan dokumentasi (Setyawan, 2018).

Tabel 2.6 Simbol *Flowchart* (Setyawan, 2018)

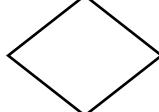
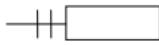
Simbol	Nama	Funi
	Terminal	Menunjukkan sebuah awal dan akhir sebuah proses
	Garis Alir	Menunjukkan arah flow algoritma, dari sebuah proses ke proses berikutnya
	Proses	Menyatakan pengolahan data menjadi data baru (output)
	Keputusan (Decision)	Dalam memproses data juga perlu adanya sebuah keputusan atau kondisi tertentu. Decision ini mempunyai 2 tipe keluaran digunakan melanjutkan jalanya proses dengan kondisi yang berbeda
	Input dan Output	Menyatakan bahwa digunakan untuk memasukan sebuah data dan keluaran data. begitu sebaliknya

Simbol (Lanjutan)	Nama (Lanjutan)	Funsi (Lanjutan)
	Predefined Process	Digunakan untuk menunjukan sebuah suatu proses yang kompleks
	Inialisasi	Menyatakan sebuah operasi dan tidak memiliki efek khusus selain mempersiapkan untuk sebuah nilai
	Konektor Dalam Halaman	Menyatakan sebuah hubungan suatu Langkah ke Langkah yang lainya
	Konektor Luar Halaman	Penghubung digunakan untuk menghubungkan suatu Langkah ke Langkah berikutnya.
	Storage	Simbol ini untuk digunakan masukan dari disk atau output disimpan ke disk.

2.2.6.4. Entity Relationship Diagram

(Sukrianto, 2017) Model ERD sebuah kumpulan yang berisi komponen baik berupa entitas maupun kumpulan relasi, yang masing – masing memounyai *atribut – atribut* yang mepresentasikan seluruh fakta yang ditinjau langsung dilapangan. Sehingga dapat diketauhi hubungan antara entity – entity (*class*) yang ada dengan atributnya. Selain itu bisa mepresentasikan suatu hubungan yang ada dalam pengolahan data atau entitas lain. Misalnya relasi *Many to Many*, *One to One*, dan lain sebagainya.

Tabel 2.7 Simbol ERD

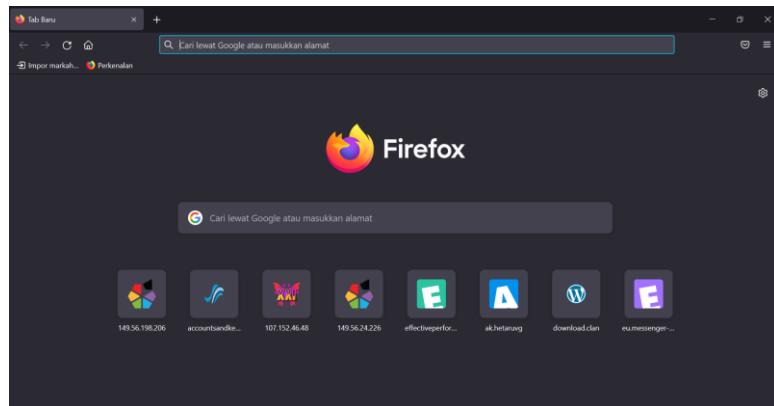
Simbol	Keterangan
	<i>Entity</i> , atau dapat dikatakan sebagai <i>class</i> berisi object yang mempunyai <i>atribut</i>
	Relasi, menyatakan sebuah hubungan (aktifitas) antara entity dengan yang lain dan saling terkait.
	Hubungan antara entity dengan derajat kardinalitas yang dinyatakan sebagai relasi <i>optional many (Many)</i>
	Hubungan antara entity dengan derajat kardinalitas yang dinyatakan sebagai relasi <i>optional alone (One)</i>
	Relasi antara entity dengan dengan derajat kardinalitas yang dinyatakan sebagai <i>mandatory many</i>
	Relasi antara entity dengan derajat kardinalitas relasi yang dinyatakan sebagai <i>mandatory one</i>

Sumber (Zefriyanto & Santoso, 2015, diacu oleh Sukrianto, 2017)

2.2.7. Perangkat lunak (Software) yang digunakan

2.2.7.1. Web Browser

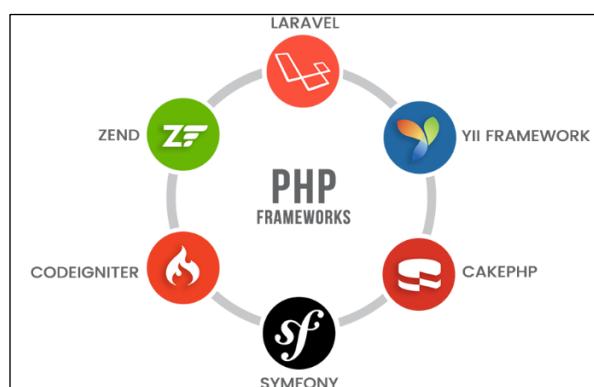
Browser salah satu sebuah perangkat lunak atau program yang digunakan untuk menjelajah internet yang melalui sebuah Komputer atau laptop. Pada saat Menjelajah sudah termasuk melintas, mengambil, dan menampilkan sebuah informasi di internet atau yang dikenal dengan *Word Wide Web*. Jadi Fungsi browser sebagai sebuah alat untuk menampilkan sumber informasi yang ada di internet kepada pengguna internet bisa berupa sebuah halaman web, gambar, video atau bagian lain dari konten website di internet (Pattianakotta et al., 2015).



Gambar 2.5 Web Browser Mozilla Firefox

2.2.7.2. PHP (Hypertext Processor)

PHP singkatan dari PHP Hypertext Processor digunakan untuk sebagai bahasa script server – side atau yang dikenal sebagai backend development web yang disisipkan pada dokumen HTML. Pengguna (user) PHP mungkin Web dapat dibuat dinamis sehingga maintenance web tersebut dapat lebih mudah dan praktis. Bahasa PHP juga Bahasa Pemrograman yang bersifat Open-Source yang disebarluaskan dan terlicensi secara gratis dan dapat di download secara bebas dari situs resminya, PHP juga ditulis dengan menggunakan Bahasa C (Suhartanto, 2012). Selain itu Bahasa PHP juga mempunyai Beberapa Framework seperti halnya *Laravel* dan *Codeigniter*.



Gambar 2.6 Framework PHP

2.2.7.3. MySQL

MySQL menjadi perangkat lunak yang digolongkan sebagai DBMS (database management system) yang bersifat open source. Kata Open source

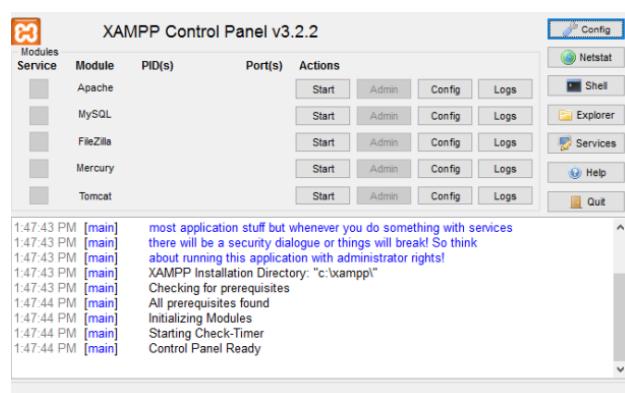
menyatakan bahwa software menyertakan code yang dipakai untuk membuat MySQL. Selain itu bentuk *executable*-nya atau kode yang dapat dijalankan atau running secara langsung dalam sistem operasi juga bisa diperoleh secara gratis dengan mendownload di internet (Nurmalina & Santoso, 2017, p. 86).



Gambar 2.7 Logo MySQL

2.2.7.4. XAMPP

XAMPP sebuah perangkat lunak yang mendukung banyak sistem atau istilah lain suatu paket perangkat lunak yang di dalamnya berisi banyak aplikasi (perangkat lunak). Selain itu Fungsi dari XAMPP sebagai server yang berdiri sendiri (localhost) terdiri dari program apache, HTTP, server MySQL database, dan penerjemah Bahasa pemrograman PHP dan Perl (Hornberger, 2002 diacu oleh (Akik & Fitria Shabrina, 2018)).



Gambar 2.8 Tampilan XAMPP

BAB III

METODE PENELITIAN

3.1. Bahan Penelitian

Bahan penelitian yang akan digunakan dalam Penelitian skripsi ini, hasil dari survey dan observasi yang telah dilakukan dan data tersebut dapat dijadikan sebagai tolak ukur atau acuan agar keakuratan dari data tersebut tetap terjaga kebenarannya dalam mendapatkan atau mengambil data tersebut. Data yang dibutuhkan untuk penelitian skripsi ini yakni berupa data PK Pendataan Keluarga 2021 di desa Semanding tepatnya di RT 01 dan RT 02 RW. V pada tahun ini yang jadikan sebagai acuan sistem untuk menghasilkan sebuah keluaran atau output. Terdapat jenis data atau bahan penelitian sebagai berikut :

1. Data Primer

Data Primer salah satu sebuah data yang diperoleh secara langsung dari sumber yang ada di lapangan, pelaku penilitian mendapatkan data primer berupa data seperti penghasilan , status bangunan , Sumber air , luas rumah , lansia , Ibu Hamil , Anak Sekolah , Jenis Dinding , Jenis Lantai , Jenis Atap , asset yang dimiliki. Data tersebut dapat dijadikan sebagai data kriteria dalam menentukan keputusan calon penerima PKH (Program Keluarga Harapan) dan data kriteria kriteria tersebut di dapatkan dari study Pustaka jurnal sebelumnya yang membahas tentang Pembagian PKH.

2. Data Sekunder

Data sekunder salah satu sebuah data yang diperoleh dari dengan cara membaca , mencari, dan mengkaji ulang materi - materi antara lain seperti berupa buku, makalah , karya ilmiah , artikel , atau dari situs internet dengan penelitian skripsi ini juga memperoleh informasi baik berupa konsep , teori maupun tutorial yang menunjang penelitian. data sekunder dapat diartikan sebagai memperoleh data yang tidak secara langsung.

3.1.1. Teknik Pengambilan data

Metode Pengambilan data yang digunakan dalam penelitian ini menggunakan salah satu tahapan proses yang ada dalam KDD (*Knowledge Discovery in Database*) yakni salah satunya tahap Selection, yang mana pemilihan data yang harus dilakukan sebelum tahap selanjutnya atau proses menggali informasi. Setelah proses penyeleksian terhadap sekumpulan data disimpan kedalam suatu berkas atau file yang terpisah.

3.1.2. Metode Pengumpulan data

Jenis Penelitian ini jenis penelitian Deskriptif Kuantitatif yang digunakan dalam penelitian ini menggunakan jenis data dan sumber data baik primer maupun sekunder .

data sekunder berupa data kriteria kriteria bobot yang dapat dari journal penelitian sebelumnya yang menunjang dalam keputusan penerima PKH. Sedangkan data primer dapat berupa data hasil observasi dan survey atau wawancara dengan seseorang yang melakukan pendataan yang terkait dengan kriteria yang ada pada setiap objek yang memiliki karakteristik yang berbeda - beda.

3.2. Metode Penelitian

Pelaku Penilitian menggunakan metode penelitian kuantitatif pada penlitian skripsi ini. Metode penelitian kuantitatif menjadi metode yang berfokus pada aspek pemahaman yang mendalam atau lebih kepada suatu masalah yang ada dari pada melihat sebuah permasalahan. Penelitian kuantitatif sebuah penlitian yang bersifat deskripsi yang mengacu pada sebuah dataset dengan memanfaatkan teori (metode) yang ada dalam penelitian teori yang digunakan yakni salah satu teori data mining sebagai bahan pendukung.

Metode kuantitatif memiliki tujuan dimana menjelaskan suatu peristiwa (fenomena) dengan lebih mendalam dengan cara mengumpulkan data yang sedalam

– dalamnya. Pada penelitian kuantitatif menunjukkan jika semakin dan detai; suatu data yang diteliti maka dapat diartikan semakin baik pula kualitas penelitian yang dilakukan.

Pada penelitian ini terdapat kerangka kerja penelitian yang pelaku penelitian yang dapat dilakukan dengan diuraikan pembahasan dari masing – masing tahap dalam penelitian yaitu sebagai berikut :

1. Pengumpulan data

Dalam tahap awal ini, Pengumpulan data dilakukan bertujuan untuk mendapatkan informasi yang dibutuhkan agar dapat mencapai tujuan penelitian. Dalam penelitian skripsi teknik mengumpulkan data antara lain yakni Observasi , Wawancara dan Study Pustaka.

2. Transformasi data

Tahap Selanjutnya, setelah pengumpulan data proses selanjutnya Merubah data atau Mentransformasikan data dimana sebuah proses perubahan data dari yang asli menjadi data yang dapat diolah dan sesuai persyaratan dengan metode perhitungan yang digunakan. Contoh dalam penelitian ini metode perhitungan data menggunakan 2 metode yaitu *Algoritma C4.5* dan *Clustering K – Means*.

3. Analisa Data

Analisa Data suatu tahapan atau kegiatan merubah hasil Penelitian (perhitungan) menjadi informasi yang dapat digunakan untuk mengambil keputusan atau menyatakan suatu kesimpulan pada penelitian tersebut.

4. Perancangan Sistem

Perancangan sistem dimana suatu kegiatan merancang dan menentukan cara mengolah sistem informasi yang akan di bangun dari hasil analisa pada sebuah sistem yang akan di bangun. Sehingga sistem bisa memenuhi kebutuhan user termasuk diantaranya baik merancang UML dan User Interface atau Prototype.

5. Pembangunan Sistem

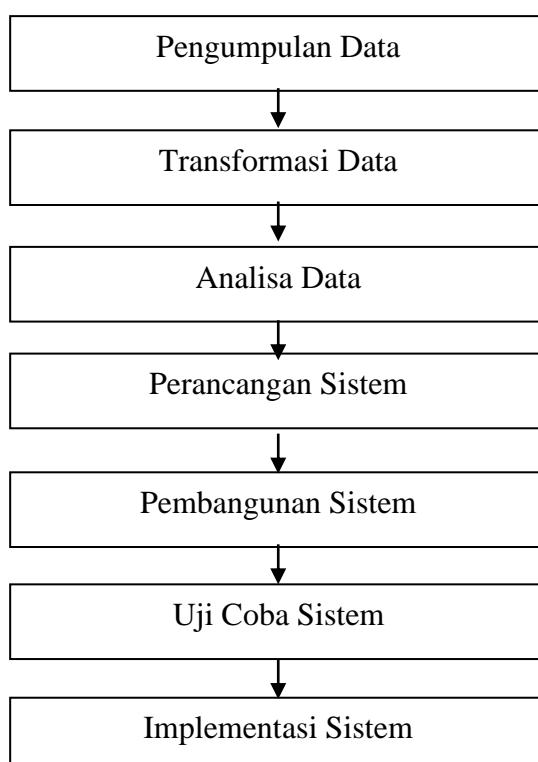
Dalam Tahap ini , sebuah Pengimplementasian dari Perancangan Sistem dimana Membangun sebuah sistem informasi yang awalnya berupa konsep atau rancangan yang dapat berupa prototype menjadi sebuah sistem yang utuh nantinya atau dapat dikatakan tahap (kegiatan) mengcoding.

6. Uji Coba Sistem

Testing sistem suatu proses (kegiatan) yang melakukan penilaian terhadap sistem yang sudah di bangun, apakah sistem yang di bangun tersebut sudah sesuai yang dinginkan selain itu Testing dapat dijadikan sebuah acuan evaluasi terhadap sistem yang di bangun apakah sistem mempunyai kelemahan atau kelebihan dari sistem yang di bangun.

7. Implementasi Sistem

Implementasi Sistem suatu tahapan yang dilakukan menyelesaikan desain sistem (sistem yang dibangun) yang telah disetujui yang sebelumnya sudah melakukan proses Testing dan di terapkan kepada tujuan awal untuk apa nantinya sistem digunakan.



Gambar 3.1 Kerangka Kerja Penelitian

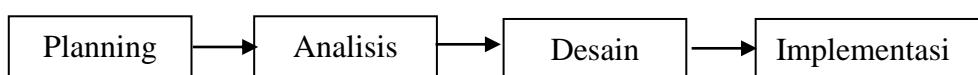
3.3. Metodologi Pengembangan Sistem

Menurut Azhar Susanto yang diacu oleh (Setiya Budi et al., 2016) mengatakan bahwa SDLC (System Development Life Cycle) menjadi salah satu metode pengembangan Perangkat Lunak (Sistem Informasi) yang populer saat Perangkat Lunak yang pertama kali dibuat.

(Setiya Budi et al., 2016) menyatakan bahwa Terdapat empat Tahapan (Langkah) yang mana membangun atau mengembangkan sebuah sistem informasi dengan menggunakan SDLC yaitu antara lain : planning , analysis, design, dan implementation. Selain itu dalam implementasi SDLC mempunyai berbagai metodologi yang dapat dipergunakan. Penggunaan metodologi akan berbagai macam metodologi tergantung kepada penekanannya, apakah terhadap bisnis proses atau pada data pendukung bisnis.

Berdasarkan Pengertian dan penjelasan secara umum dapat dinyatakan bahwa proses pengembangan perangkat mengikuti tahap – tahap sebagai berikut (Budi dkk, 2016):

1. Menentukan APA yang harus dikerjakan oleh perangkat lunak dalam waktu tertentu
2. Mendefinisikan suatu Perangkat Lunak BAGAIMANA dibuat yang mencakup arsitektur perangkat lunaknya, user interface , algoritma dan lain sebagainya.
3. Penerapan (menulis coding) dan serta pengujian unit – unit mendebug suatu program.
4. Validasi perangkat lunak (Software) secara keseluruhan (Pengujian Sistem)



Gambar 3.2 System Development Life Cycle

3.4. Alat Penunjang Penelitian

Dalam Penelitian tentunya ada alat (tools) untuk mendukung jalanya Penelitian ini. Alat penunjang Penelitian dapat dikategorikan diantaranya yakni

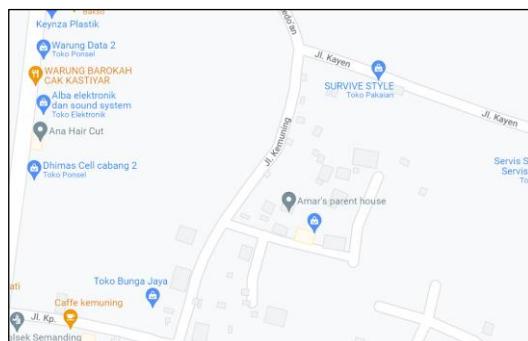
antara lain Perangkat Keras dan Perangkat Lunak Software, sebagai analisis kekurangan dan kebutuhan yang harus dipenuhi baik dalam proses perancangan sistem maupun proses penelitian nantinya.

Tabel 3.1 Daftar Alat penunjang Penelitian

No	Kategori	Kebutuhan
1	Perangkat Lunak	Windows 11 Home Single Languange 64-bit
		Program Aplikasi Text Editor (Coding) Visual Studio Code
		Program Aplikasi database MYSQL
		Program paket Aplikasi server yaitu XAMPP
		Web Browser (Google Chrome dll)
		MS. Word 2019 dan MS. Excel 2019
2	Perangkat keras	Laptop Lenovo Ideapad 3 core I5 gen 10 8gb RAM 2 GB 512GB SSD
		Printer
		Kertas (A4, F4, Folio)
		Alat Tulis (Bolpen dan lain Sebagainya)

3.5. Tempat Penelitian

Penelitian ini dilakukan bertempat di Dsn. Semanding Timur, Ds. Semanding RT 01 dan RT 02 RW V Kec. Semanding Kab. Tuban. Lokasi Tersebut dipilih karena memiliki beberapa aspek pendukung yang diperlukan agar Penelitian yang dilakukan dapat berjalan dengan baik dan lancar.



Gambar 3.3 Lokasi Ds. Semanding Timur RW 05

BAB IV

JADWAL PENELITIAN

No	Kegiatan	Bulan Pelaksanaan Penelitian																															
		Oktober 2021				November 2021				Desember 2021				Januari 2022				Februari 2022				Maret 2022				April 2022				Mei 2022			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Pengajuan Judul																																
2	Penyusunan Skripsi																																
3	Bimbingan Skripsi																																
4	Seminar Proposal																																
5	Penyusunan Skripsi Lanjutan																																
6	Bimbingan Skripsi Lanjutan																																
7	Sidang Skripsi																																

DAFTAR PUSTAKA

- Adisanjaya Suleman, S., & Resnawaty, R. (2017). Program Keluarga Harapan (PKH) Antara Perlindungan Sosial dan Pengetasan Kemiskinan. *PROSIDING KS: RISET & PKM*, 4(1), 1–140.
- Akik, H., & Fitria Shabrina, A. (2018). JURNAL MANAJEMEN INFORMATIKA. *JUMIKA*, 5(2).
- Ardiansyah Sembiring, M., Fitri Larasati Sibuea, M., Sapta, A., Studi Sistem Informasi, P., & Royal, S. (2018). ANALISA KINERJA ALGORITMA C.45 DALAM MEMPREDIKSI HASIL BELAJAR. *Journal of Science and Social Research*, 1, 73–79. <http://jurnal.goretanpena.com/index.php/JSSR>
- Fatmawati, & Munajat, J. (2018). IMPLEMENTASI MODEL WATERFALL PADA SISTEM INFORMASI PERSEDIAAN BARANG BERBASIS WEB (STUDI KASUS: PT.PAMINDO TIGA T). *MEDIA INFORMATIKA BUDIDARMA*, 2(2), 1–9. <http://ejurnal.stmik-budidarma.ac.id/index.php/mib>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases) (© AAAI). *AI Magazine*, 17, 37–54. www.aaai.org/Pubs/Magazines/AIM/1996/17-01/17-01-0003.pdf
- Heriyanto, Y. (2018). Perancangan Sistem Informasi Rental Mobil Berbasis Web Pada PT.APM Rent Car. *Jurnal Intra-Tech*, 2(2), 64–77.
- Nasrullah, A. H. (2018). ILKOM Jurnal Ilmiah Volume 10 Nomor 2 Agustus 2018. *ILKOM Jurnal Ilmiah*, 10(2), 244–250.
- Nofriansyah, D., & Nurcahyo, W. G. (2019). *ALGORITMA DATA MINING DAN PENGUJIAN* (M. C. Sartono, Ed.). Deepublisher.
- Nurmalina, R., & Santoso. (2017). Perencanaan dan Pengembangan Aplikasi Absensi Mahasiswa Menggunakan Smart Card Guna Pengembangan Kampus Cerdas (Studi Kasus Politeknik Negeri Tanah Laut). *Jurnal Integrasi*, 9(1), 84–91.
- Pattianakotta, A., A.E. Snsuw, A., & S.M. Lumenta, A. (2015). Sistem Informasi Arsip Dolumen Kantor Pelayanan Kekayaan Negara Dan Lelang manado. *E-Journal Teknik Elektro Dan Komputer*, 4(7), 8–14.
- Putri Pratiwi, I., Ferdinandus, F., Daniel Limantara, A., Tinggi Teknologi Cahaya Surya Kediri, S., & Tinggi Teknik Surabaya, S. (2019). Sistem Pendukung Keputusan Penerima Program Keluarga Harapan (PKH) Menggunakan Metode Simple Additive Weighting. *Jurnal Teknik Informatika, Sistem Informasi , Dan Ilmu Komputer*, 8(2), 2580–2399.
- Putung, K. D., Lumenta, A., & Jacobus, A. (2016). PENERAPAN SISTEM TEMU KEMBALI INFORMASI PADA KUMPULAN DOKUMEN SKRIPSI. *18 E-Journal Teknik Informatika*, 8(1).
- Randa, D. D. (2018). PERANCANGAN DAN PENERAPAN SISTEM INVENTORY BARANG PADA TOKO BIG STORE PADANG DENGAN MENGGUNAKAN BAHASA PEMROGRAMAN JAVA DAN DATABASE MYSQL. *INTECOMS*:

Journal of Information Technology and Computer Science, 1(2), 224–230.
<https://doi.org/10.31539/intecoms.v1i2.294>

- Rosa dkk. (2013). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek* (Vol. 1). Informatikla. http://pustaka.fti.unand.ac.id//index.php?p=show_detail&id=535
- Rufiyanto, A., & dkk. (2021). *PENERAPAN ALGORITMA C4.5 UNTUK PREKDISI KEPUASAN MAHASISWA* (A. Dzatin Nabila, Ed.). Deepublisher.
- Saragih, I. A. (2020). *PENERAPAN METODE C4.5 UNTUK PENENTUAN KELAYAKAN PENERIMA PROGRAM KELUARGA HARAPAN*.
- Setiya Budi, D., Azhima Yoga Siswa, T., & Abijono, H. (2016). Analisis Pemilihan Penerapan Proyek Metodologi Pengembangan Rekayasa Perangkat Lunak. *24 TEKNIKA*, 5(1), 24–31.
- Setyawan, U. (2018). *SISTEM INFORMASI GEOGRAFIS PEMETAAN WARGA MISKIN YANG PANTAS MENDAPATKAN BANTUAN SOSIAL MENGGUNAKAN METODE CLUSTERING K - MEANS*.
- Suhartanto, M. (2012). Pembuatan Website Sekolah Menengah Pertama Negeri 3 Delanggu. *Journal Speed-Sentra Penelitian Engineering Dan Edukasi*, 4(1), 1–8. www.oreilly.com
- Sukrianto, D. (2017). Penerapan Teknologi Barcode Pada Pengolahan Data Pembayaran Sumbangan Pembinaan Pendidikan (SPP). *Jurnal Intra-Tech*, 1(2), 18–27.
- Suyanto. (2019). *DATA MINING UNTUK KLASIFIKASI DAN KLASTERISASI DATA* (Vol. 1). Informatika Bandung.
<https://opac.perpusnas.go.id/DetailOpac.aspx?id=1059041>
- Tahyudin, I. (2014). *SISTEM PENDUKUNG KEPUTUSAN (SPK) KONSEP DASAR DAN PENERAPANNYA DAN DATA MINING* (Edisi Pertama, Vol. 1).
https://www.google.co.id/books/edition/SISTEM_PENDUKUNG_KEPUTUSAN_SPK_KONSEP_DA/kFZFEAAAQBAJ?hl=id&gbpv=0
- Urva, G., & Fauzi Siregar, H. (2015). Pemodelan UML E-Marketing Minyak Goreng. *JURTEKSI ROYAL*, 2, 92–101.
- Wanto, A., Siregar, H. N. M., Windarto, P. A., Hartama, D., Ginantra, R., Napitupulu, D., Negara, S. E., Lubis, R. M., Dewi, V. S., & Prianto, C. (2020). *DATA MINING ALGORITMA DAN IMPLEMENTASI* (T. Limbong, Ed.; Pertama, Vol. 1). Yayasan Kita Menulis.
- Wijaya, A., Arifin, M., & Soebijono, T. (2013). SISTEM INFORMASI PERENCANAAN PERSEDIAAN BARANG. *JSIKA* , 2, 14–20.
<http://jurnal.stikom.edu/index.php/jsika>