

BAB II

KAJIAN PUSTAKA

2.1 Knowledge Discovery In Database (KDD)

Knowledge Discovery in Database (KDD) adalah merupakan metode dalam mendapatkan keahlian basis data yang ada. Dalam *database* diperoleh grafik-grafik yang saling terhubung satu sama lainnya. Pengambilan keputusan dari hasil keahlian yang didapatkan melalui tahapan yang berguna sebagai basis pengetahuan (*knowledge base*). Istilah *Knowledge Discovery in Database (KDD)* dan *data mining* biasanya dimanfaatkan secara bergantian dalam memperjelas tahapan penggalian informasi tertanam pada *database* besar. Dari kedua perumpamaan tersebut mempunyai ide tidak serupa namun memiliki keterikatan diantaranya, dan satu diantaranya yaitu bagian dari keseluruhan prosedur *Knowledge Discovery in Database (KDD)* adalah *data mining*. (Mardi, 2018)

Tahapan *Knowledge Discovery in Database (KDD)* secara ilustrasi bisa dijelaskan sebagai berikut:

1. Data Selection

Sebelum proses penggalian informasi *Knowledge Discovery in Database (KDD)* dimulai, maka diperlukan proses penyeleksian data dari sekelompok data operasional. Hasil dari penyeleksian data tersebut disimpan kedalam berkas terpisah dari database operasional untuk proses data *mining*.

2. *Pre-processing / Cleaning*

Tahapan pembersihan data pada saat fokus *Knowledge Discovery in Database* (KDD) dilaksanakan sebelum proses data *mining* dimulai. Tahapan penyortiran data meliputi yaitu penyortiran data ganda, mendalami inkonsisten data, dan mengubah data yang salah cetak. Selain itu juga ada proses *enrichment*, adalah data atau informasi luar yang dibutuhkan oleh *Knowledge Discovery in Database* (KDD) yang valid melalui tahapan memperbanyak data yang telah diketahui sebelumnya.

3. *Transformation Coding*

Tahapan *coding* dalam *Knowledge Discovery in Database* (KDD) ialah prosedur yang bergantung pada macam atau bentuk dari informasi yang akan dicari pada *database* serta prosedur transmudasi atas berita yang sudah diambil berdasarkan kemauan pada data *mining*.

4. *Data Mining*

Data *mining* merupakan suatu cara dalam mencari sketsa serta berita menarik dalam yang telah terpilih oleh teknik dan metode tertentu. Data *mining* memiliki banyak teknik dan metode atau *algoritma* yang beranekragam sehingga pemilihan metode dan *algoritma* yang tepat sangat bergantung pada hasil dan tahapan dari *Knowledge Discovery in Database* (KDD) secara keseluruhan.

5. *Interpretation / Evaluation*

Bentuk informasi dari proses data mining yang telah diperoleh yaitu harus memiliki bentuk yang sederhana yang gampang dipahami oleh pihak yang bersangkutan. Proses ini adalah bagian dari *Knowledge Discovery in Database*

(KDD) yang diartikan sebagai *interpretation*. proses ini mencakup pemeriksaan apakah sketsa atau berita yang didapatkan bertolak belakang dengan fakta atau hipotesis sebelumnya yang sudah ada. (Mardi, 2018)

2.2 Data Mining

Data *mining* memanfaatkan prosedur khusus untuk memperoleh skema maupun berita merampas atas sebuah berita. Tahapan dari *Data mining* yang digunakan dalam mengekstrak serta mengidentifikasi suatu informasi yang dibutuhkan oleh *database* besar dengan cara teknik statistik, matematika, kecerdasan buatan, dan *machine*. (Mardi, 2018)

Tahapan induk dari kumpulan informasi yang dimana didalamnya terlibat beberapa bidang yaitu bidang bidang ilmu statistik, mesin pembelajaran, dan sistem *managemen database* dalam pemanfaatan *algoritma* serta dipergunakan untuk menghasilkan data tersembunyi dalam suatu *database* yang besar merupakan pengertian dari data *mining*. (Yanto & Khoiriah, 2015)

Data *mining* merupakan sebuah cara yang dibantu oleh metode statistik matematika, kecerdasan buatan, dan *machine learning* untuk menghasilkan dan membuktikan berita yang bermanfaat serta ilmu yang berhubungan dengan *database* besar. Data *mining* juga dapat diartikan sebagai rangkaian cara pada proses penggalian data yang besar serta nilai tambah suatu kelompok data yang merupakan ilmu yang tidak diketahui secara manual selama ini. (Faisal, 2019)

2.2.1 Manfaat Data Mining

Dalam data *mining* terdapat 2 dimensi, yaitu dimensi komersial dan dimensi keilmuan. (Vulandari, 2017) yaitu sebagai berikut:

1. Dari dimensi komersial, penggunaan data *mining* dapat peruntukkan menanggapi meluapnya isi data, dengan cara memanfaatkan metode komputasi sehingga dapat memperoleh berita-berita yang diperlukan aset supaya dapat menaikkan daya saing suatu kelompok.
2. Dari dimensi keilmuan, *data mining* dimanfaatkan dalam menggambar, menganalisis serta mengamankan besar yang berkarakter *realtime*.

2.2.2 Fungsi Data Mining

Manfaat-manfaat yang sering dipergunakan pada *data mining* menurut haskett 2000 (Vulandari, 2017), sebagai berikut:

- 1) *Assosiation*, merupakan tahapan yang digunakan dalam memperoleh kebijakan asosiasi kombinasi barang pada suatu waktu.
- 2) *Sequence*, merupakan tahapan yang digunakan untuk menenukan kebijakan asosiasi kombinasi barang dalam waktu banyak.
- 3) *Clustering*, merupakan tahapan dari kumpulan data serta memiliki kumpulan *database* sama.
- 4) *Classification*, merupakan tahapan yang digunakan untuk membedakan struktur maupun kelas data yang bertujuan supaya objek tersembunyi.

- 5) *Regression*, merupakan tahapan pembagian berita marka perkiraan.
- 6) *Forecasting*, merupakan pembedaan nilai perkiraan dari sekelompok bentuk.
- 7) *Solution*, merupakan penemuan biang masalah yang dihadapi untuk pemugutan kesimpulan.

2.2.3 Penerapan *Data Mining*

Menurut (Vulandari, 2017), penerapan data *mining* terdiri dari 3 penerapan yaitu sebagai berikut:

1. Analisa Pasar Dan Manajemen

Asal mula berita atas diaplikasikan seolah negosiasi nota utang, nota kelompok *group* eksklusif, nota potongan, protes konsumen, ditingkatkan via belajar mengenai konsep tumbuh masa. Macam-macam jalan keluar yang bisa diatasi oleh data *mining* antara lain:

- a) Menghabisi sasaran market
 - b) Melihat kondisi daya beli dimasa depan
 - c) Analisis lintas pasar
 - d) tentang *costumer*
 - e) Melihat keperluan *costumer*
 - f) Mengapresiasi kelayakan *custumer*
 - g) Berita pendapatan
- #### 2. Analisa Perusahaan Dan Manajemen Resiko

- a) Perencanaan keuangan dan evaluasi aset
- b) Perencanaan sumber daya
- c) Persaingan

3. Telekomunikasi

Data *mining* digunakan untuk meningkatkan pelayanan langsung dengan banyaknya penjualan yang diterima serta menyortir penjualan yang harus dilayani secara manual.

2.2.4 Teknik Pembelajaran *Data Mining*

Terdapat tiga metode utama yang merupakan bagian dari metode data mining yang sangat erat dengan penemuan (*discovery*) dan pembelajaran (*learning*) yaitu :

1. *Supervised Learning*

Supervised learning adalah metode yang sering dipergunakan. Metode ini juga disebut "*Programming By Example*". Metode ini sangat bergantung tahapan latihan pada data historis yang memiliki bagian-bagian pemetaan terhadap hasil pengolahan *algoritma data mining*.

2. *Unsupervised Learning*

Metode ini kebalikan dari metode *supervised learning*. Metode ini sangat bergantung kepada pemanfaatan *algoritma* yang mendeteksi semua bentuk, seperti *assosiation* dan *sequence*, yang ada pada spesifik dalam data pemasukan.

3. *Reinforcement Learning*

Metode ini memiliki *control* adaptif namun paling jarang digunakan.

2.2.5 Kategori *Data Mining*

(Vulandari, 2017) *Data mining* terdiri dari 2 kategori utama yaitu seperti berikut:

1. Prediktif

Berfungsi dalam menentukan nilai atribut tertentu yang dilihat dari atribut-atribut lainnya.

2. Deskriptif

Berfungsi dalam menentukan bentuk maupun struktur dalam merangkum hubungan utama dalam data.

2.3 Metode *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (Mardi, 2018), yaitu:

1. ***Description (Deskripsi)***

Deskripsi merupakan bentuk dari kecenderungan yang sering memberikan kemungkinan penjelasan dari suatu pola.

2. ***Estimation (Estimasi)***

Estimasi merupakan hasil dari variabel target berdasarkan hasil prediksi yang hampir sama fungsinya dengan klasifikasi.

3. ***Prediction (Prediksi)***

Prediksi ini sendiri memiliki kemiripan dengan klasifikasi dan estimasi, tetapi prediksi lebih mengarah ke hasil dimasa yang akan datang.

4. *Classification* (Klasifikasi)

Merupakan bagaian dari kategori yang dikelompokkan menjadi beberapa bagian.

5. *Clustering* (Pengkusteran)

Merupakan kumpulan objek-objek yang hampir sama satu dengan yang lainnnya serta memiliki ketidaksamaan dengan objek dri kelompok lainya

6. *Association* (Asosiasi)

Merukan bagian dari data mining yang dari menemukan item yang muncul dalam watu yang bersamaan.

Berdasarkan penjelasan dari beberapa kelompok yng telah dijelaskan diatas, maka metode dari teknik data mining yang akan digunakan dalam penelitian ini yaitu menggunakan teknik *Classification* (Klasifikasi).

2.3.1 *Classification* (Klasifikasi)

Proses pengklasifikasian merupakan salah satu pekerjaan dari data *mining*. Kemunculan klasifikasi awalnya dilakukan pada bagian tanaman yang dilakukan pada bentuk tertentu, seperti (*Carolus Linnaeus*) yang dilakukan oleh *Carolus von Linne* yang mulanya tidak diketahui nama pada saat pengklasifikasian bentuk berdasarkan karakter fisik. Selanjutnya dia dikenal sebagai bapak klasifikasi. Beberapa metode klasifikasi umum digunakan dalam data *mining* adalah: Pohon

keputusan, Pengklasifikasi bayes/*naive bayes*, Jaringan saraf tiruan, Analisis statistik, *Algoritma* genetik, *Rough sets*, Pengklasifikasi *k-nearest neighbour*, Metode berbasis aturan, *Memory based reasoning*, *Support vector machine*. (Faisal, 2019)

Klasifikasi data merupakan proses pengklasifikasian dalam ruang yang berbeda-beda yang ditetapkan sebagai model klasifikasi dari sebuah *database* yang berupa pola-pola yang ditemukan oleh himpunan obyek. Klasifikasi itu sendiri memiliki tujuan yaitu pola dari *training set* kedalam kelompok atau kelas yang sesuai berdasarkan atribut yang berbeda-beda, pola tersebut selanjutnya digunakan untuk pengklasifikasian atribut dengan kelas yang belum diketahui sebelumnya. (Azwanti, 2018)

2.3.2 Pohon Keputusan

Pohon keputusan dikenal sebagai contoh bentuk pohon (*tree*) yang mana setiap *node* mencontohkan atribut, cabangnya mencontohkan nilai dari atribut dan daun mencontohkan kelas. *Root* merupakan pohon keputusan paling atas dari sebuah *node*. Pohon keputusan memiliki keuntungan dalam mengekstrak dan mencontohkan bentuk dalam klasifikasi aturan *if-then* salah satunya adalah C4.5 (Azwanti, 2018)

Pohon keputusan atau *decission tree* adalah beberapa metode dari teknik data maning yang dimanfaatkan sebagai pengklasifikasian. Pohon keputusan adalah aturan yang sangat gampang dipelajari dengan bahasa sehari-hari serta dapat merubah kebenaran yang besar menjadi sebuah merepresentasikan aturan.

Sebuah pohon keputusan adalah sebuah konsep yang dapat dimanfaatkan dalam proses pembagian elemen-elemen data besar menjadi sebuah tumpukan *record* yang berukuran mini yang diterapkan sebagai serangkaian aturan keputusan. Salah satu *algoritma* yang dimanfaatkan untuk pembentukan pohon keputusan yaitu *algoritma C4.5*. (Faisal, 2019)

Pola dari pohon kesimpulan sengaja disamakan dengan arsitektur yang ada pada sebuah pohon dan penegasannya, ialah seperti berikut ini:

1. Simpul bagian akar

Merupakan bagian akar paling atas dari pola pada pohon keputusan.

2. Simpul internal

Merupakan bagian cabang lanjutan dari simpul akar. Simpul ini akan mengeluarkan simpul maksimal 2 dari 1 simpul yang dimasukkan.

3. Simpul daun

Merupakan bagian paling ujung dari semua pohon yang tidak memiliki turunan atau keluaran.

Struktur pokok kesimpulan memiliki tiga macam *node* yaitu (Elisa & Ervina, 2018) :

1. *Rood Node* yaitu *node* pertama adalah *node* yang mempunyai *input* namun

tidak mempunyai *output* tetapi banyak *output* juga.

2. *Internal Node* yaitu *node* pembagian adalah *node* dengan 1 *input* dan *output*

yang banyak.

3. *Leaf Node* yaitu *node* akhir adalah *node* dengan *input* saja.

2.3.3 Algoritma C4.5

Salah satu algoritma dalam membentuk suatu pokok kesimpulan yang biasa dipakai adalah *algoritma C4.5*. *Algoritma C4.5* adalah pengembangan atas ID3, pengembangannya untuk melumpuhkan nilai yang hilang, biasanya melumpuhkan data selanjutnya serta harapan. *Algoritma C4.5* merupakan metode untuk menstruktur sebuah pohon keputusan berkenaan dengan berita yang didapatkan terlebih dahulu. Gaya pertama yang melandasi *algoritma* ini adalah pengstruktur potoh kesimpulan berkenaan gaya cabang yang memiliki marka *gain* tertinggi berkenaan dengan marka *entropy* cabang tersebut semisal *under step* dari *algoritma* ini yaitu mempunyai nilai *gain* tertinggi berdasarkan nilai *entropy* bagian dari proses pengklasifikasian. Rumus *gain* tertinggi dapat dilihat sebagai berikut ini (Elisa & Ervina, 2018).

1) Perhitungan *Gain* Tertinggi

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

Rumus 2. 1 Perhitungan *Gain*

Dimana:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Selanjutnya melakukan perhitungan nilai *entropy* dapat disimak pada persamaan berikut ini:

$$Entropy (S) = \sum_{i=1}^n - P_i * \log_2 P_i$$

Rumus 2. 2 Perhitungan *Entropy*

Dimana:

S : Himpunan kasus

A : Fitur

N : Jumlah partisi *S*

P_i : Proporsi dari *S_i* terhadap *S*

2.4 Software Pendukung

2.4.1 *Waikato Environment for Knowledge Analysis (WEKA)*

Waikato Environment for Knowledge Analysis adalah sebuah kelompok alat pembelajaran mesin praktis yang disingkat WEKA. Aplikasi ini diciptakan di Universitas Waikato, New Zealand bagi ulasan, keahlian dan macam-macam aplikasi. Keuntungan menggunakan WEKA yaitu khususnya klasifikasi yang dilandasi oleh pendekatan-pendekatan sangat mampu menyelesaikan masalah-masalah data *mining* dalam dunia nyata. Selain itu WEKA juga sangat gampang diterapkan di beberapa tahapan berbeda dan WEKA menyediakan alat untuk *pre-processing* data, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. Ada beberapa metode dari data mining yang bisa dieksekusi menggunakan

aplikasi WEKA yang sering digunakan yaitu pohon keputusan (*decision tree*). Konstruksi dari pohon keputusan yang melibatkan *node* keputusan yang akan disambungkan ke setiap bagian-bagian mulai simpul akar sampai ke *node* daun (akhir) merupakan bagian yang menarik dari metode klasifikasi. Pada *node* kesimpulan indikasi bakal dites, lalu tiap-tiap reaksi atas memperoleh bagian. Tiap-tiap bagian akan ditunjukkan ke *node* lain atau ke *node* akhir untuk mendapatkan hasil yaitu berupa keputusan. (Pujiono, Amborowati, & Suyanto, 2013)

2.5 Penelitian Terdahulu

Penelitian terdahulu merupakan acuan dalam sebuah penelitian. Peneliti juga menemukan beberapa hasil penelitian terdahulu yang berhubungan dengan *algoritma* C4.5 antara lain sebagai berikut:

1. (Mardi, 2018) dalam jurnal yang berjudul **“Data Mining : Klasifikasi Menggunakan Algoritma C4.5”** menyatakan bahwa *Knowledge Discovery in Database* (KDD) merupakan bagian tertinggi dari data *mining*, dimana data dengan jumlah yang sangat besar dapat diolah kedalam data mining dengan melakukan berbagai metode untuk memperoleh sebuah berita seperti pengklasifikasian maupun memprediksi. Pohon keputusan yang dihasilkan melalui *algoritma* C4.5 agar gampang dimengerti adalah bagian dari klasifikasi dalam data *mining*.
2. (Novianti & Bahri, Tedy Rismawan, 2016) dalam jurnal berjudul **“Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan**

Siswa (Studi Kasus: Sma Negeri 1 Pontianak)” menyatakan bahwa klasifikasi penjurusan siswa berdasarkan penggolongan siswa berdasarkan kemampuan (nilai), bakat serta minat yang hampir sama sehingga pembelajaran yang akan diajarkan akan lebih fokus dan teratur. Salah satu institusi pendidikan yang berada di kota Pontianak yaitu SMA Negeri 1 dimana sekolah tersebut memiliki dua penjurusan yaitu IPA dan IPS. Kurikulum 2013 saat ini yang mengatur proses penyeleksian berdasarkan nilai rata-rata rapor SMP, nilai UAS SMP serta nilai tes MTK, IPA dan IPS. Metode algoritma C4.5 dimanfaatkan sebagai proses dalam pengklasifikasian data sehingga mampu menghasilkan pohon keputusan berdasarkan proses penyeleksian data sebelumnya sehingga menghasilkan tingkat akurasi sebesar 89.74%.

3. (Mirqotussa'adaha, Much Aziz Muslima & Budi Prasetyoa, 2017) dalam jurnal yang berjudul **“Penerapan Dizcretization dan Teknik *Bagging* Untuk Meningkatkan Akurasi Klasifikasi Berbasis *Ensemble* pada *Algoritma C4.5* dalam Mendiagnosa Diabetes”** menyatakan bagi aspek bakal kesehatan, data mining bakal difungsikan dalam memperkirakan suatu penyakit dari data rekam medis pasien. Sehingga digunakan untuk meningkatkan ketetapan *algoritma* pengkelompokkan dalam merancang berbagai *classifier* dari data *training* dalam mematenkan *discretization* dan teknik *bagging* untuk klasifikasi berbasis *ensemble* pada *algoritma C4.5* dapat meningkatkan akurasi sebesar 6,26%. Hasil ketetapan dari 68,61%, kemudian ditetapkan *discretization* dan metode *bagging* menjadi 74,87%.

4. (Purwadi, 2018) dalam jurnal yang berjudul **“Implementasi Data Mining Untuk Memprediksi Pola Pembelian Sepeda Motor Pada Showroom Cv. Viva Mas Motors Dengan Metode Algoritma C4.5”** menyatakan bahwa kecenderungan kelakuan konsumen dalam memilih sepeda motor menjadi bagian yang perlu di prediksi dengan menganalisa produksi sepeda motor yang semakin kompetitif. Fakta menyatakan bahwa permasalahan pola pembelian terhadap sepeda motor merupakan bagian yang akan diolah. metode *Algoritma C4.5* bagian dari data mining ini digunakan untuk memprediksi pola pembelian sepeda motor. *field merk*, tahun, dan harga merupakan variabel yang mempengaruhi pemilihan sepeda motor sehingga variabel ini akan dihitung nilai entropy serta gain terbesar sehingga dapat menghasilkan *root* dari *tree*.
5. (Nurul Aswanti, 2018) dalam jurnal yang berjudul **“Penerapan Algoritma C4.5 Untuk Memprediksi Besarnya Penggunaan Listrik Rumah Tangga Di Kota Batam”** menyatakan bahwa Listrik sudah menjadi kebutuhan manusia dalam kegiatan sehari-hari. Penggunaan listrik mulai dari peralatan rumah tangga seperti kulkas, kipas angin, mesin cuci hingga alat komunikasi seperti handphone, laptop dan berbagai macam alat elektronik lainnya yang sangat membutuhkan energi listrik. Batam merupakan salah satu kota besar dan padat di Indonesia salah satunya daerah Batam Center. Akibat kepadatan penduduk tersebut menyebabkan kebutuhan energi listrik meningkat. PT PLN Batam (*b'right*) merupakan perusahaan di kota batam yang menyediakan sumber energi listrik. Dalam

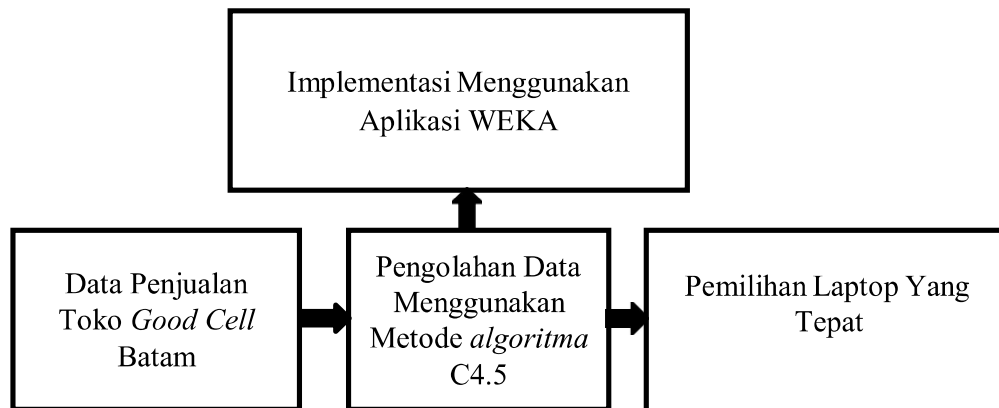
kasus ini sangat diperlukan suatu teknik yang bisa digunakan untuk memprediksi penggunaan listrik rumah tangga supaya program pemerintah dalam upaya penghematan energi listrik dapat tercapai. Salah satu teknik yang dapat digunakan dalam kasus ini yaitu teknik data *mining* dengan melakukan klasifikasi menggunakan metode algoritma C45. Hasil yang dicapai dari metode algoritma C45 berupa pohon keputusan dimana variabel jumlah anggota keluarga, luas bangunan rumah dan lama waktu dirumah menjadi variabel penting dari prediksi besarnya pemakaian listrik. Pengujian dilakukan menggunakan Weka yang menghasilkan *rule* yang sama dengan pohon keputusan yang terbentuk.

6. (Hilda Amalia, 2017) dalam jurnal yang berjudul **“Penentuan Proses Persalinan Ibu Melahirkan Menggunakan *Algoritma C4.5*”** menyatakan bahwa melahirkan merupakan sebuah karunia yang dirasakan oleh semua ibu. Proses persalinan ada dua macam yaitu normal dan operasi keduanya memiliki tujuan yang sama yaitu keselamatan ibu dan anak. Oleh sebab itu diperlukan suatu teknik dari data *mining*, sehingga dapat menghasilkan prediksi proses persalinan ibu melahirkan. *Algoritma C4.5* merupakan salah satu teknik data *mining* yang mampu menghasilkan pohon keputusan. Dengan adanya prediksi proses persalinan ini dapat dipilih proses melahirkan yang tepat sehingga dapat terhindar dari resiko persalinan. Dalam penelitian ini akan dilakukan pengolahan data-data kesehatan ibu melahirkan dengan algoritma c4.5 diperoleh hasil akurasi sebesar 90%.

7. (Kurniawan, Wicaksono, & Astuti, 2016) dalam jurnal yang berjudul **“Pemanfaatan *Educational Data Mining* (Edm) Untuk Memprediksi Masa Studi Mahasiswa Menggunakan *Algoritma C4.5* (Studi Kasus: Ti-S1 Udinus)”** Menyatakan bahwa institusi pendidikan yang memiliki ketersediaan data yang banyak diharapkan dapat dimanfaatkan dengan baik dan benar. Bentuk pembelajaran mahasiswa serta hubungan antar atribut data pendidikan menjadi fokus dalam penelitian karena sangat mempengaruhi data yang besar. Penerapan metode *Algoritma C4.5* dari data mining diharapkan dapat menemukan pola klasifikasi terhadap mahasiswa dengan kelulusan tercepat dan kelulusan yang lambat serta perkiraan data yg diuji. Sehingga *Algoritma C4.5* mampu melakukan prediksi dengan kebenaran yang baik (73,68%) terhadap masa studi mahasiswa dengan kelulusan tepat waktu dan kelulusan lambat.

2.6 Kerangka Pemikiran

Struktur pemikiran merupakan suatu struktur yang menegaskan mengenai jalan nalar yang nantinya akan dipergunakan dalam penelitian. Berikut ini adalah gambaran dari struktur pemikiran dari penelitian ini:



Sumber: Data Peneliti (2021)

Data penjualan yang berasal dari hasil data rekapan penjualan laptop selama 3 tahun pada Toko *Good Cell* Batam menjadi data *input* dalam proses penelitian, kemudian diolah ke dalam metode *algoritma* C4.5, selanjutnya data yang telah diolah dengan metode *Algoritma* C4.5 diuji menggunakan aplikasi WEKA, sehingga menghasilkan *output* berupa hasil pemilihan laptop yang tepat bagi konsumen maupun calon konsumen dari Toko *Good Cell* Batam.