

BAB II

TINJAUAN PUSTAKA

2.1. Teori Umum

Bagian berisikan penjelasan tentang istilah-istilah dan penjabaran umum yang berkaitan pada pelaksanaan pengkajian.

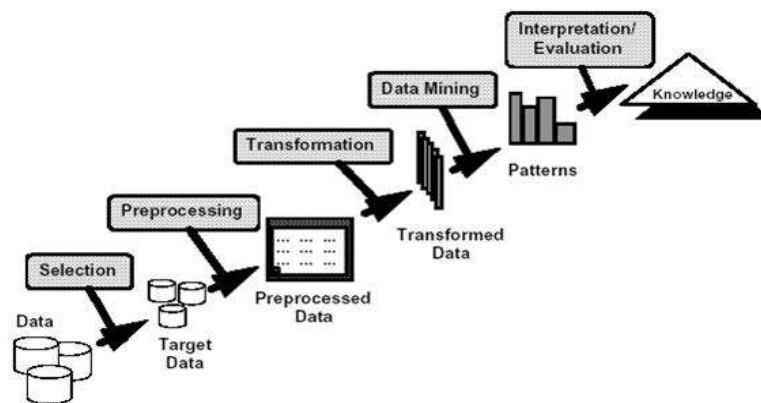
2.1.1. *Knowledge Discovery in Database (KDD)*

Berdasarkan (Marlina, 2018), Teknik untuk mengeluarkan informasi tersembunyi dari basis data yang cukup besar disebut sebagai penggalian data atau Penemuan Pengetahuan dalam Basis Data (KDD). Namun, keduanya memiliki pengertian yang berbeda namun saling berkaitan. Salah satu elemen penting dalam proses KDD secara keseluruhan adalah penggalian data. Berikut ini adalah deskripsi proses KDD secara umum:

1. Pemilihan data ialah Langkah awal dalam KDD adalah memilih data dari data aktivitas sebelum memulai tahap ekstraksi data. Data yang telah dipilih untuk data minning dimuat pada berkas yang berbeda dari data pokok operasi.
2. Preprocessing/Pembersihan merupakan Data yang digunakan sebagai subjek KDD harus dibersihkan sebelum penggalian data dapat dilakukan. Prosedur pembersihan ini melibatkan pembuangan data yang berlebihan, memastikan konsistensi data, dan memperbaiki masalah data seperti kesalahan pengetikan.
3. Mengubah data yang dipilih ke bentuk yang cocok untuk melakukan data mining dikenal sebagai transform coding. Jenis atau model data yang

didapatkan pada database mempunyai dampak yang signifikan terhadap proses kreatif pengkodean KDD.

4. Data mining ialah alur penerapan alat disebut juga pendekatan khusus agar menemukan algoritma dan data yang baik pada selection. Ada banyak pendekatan, strategi, dan algoritma data mining yang berbeda. Tujuan dan keseluruhan proses KDD memiliki peran penting dalam memilih teknik atau algoritma terbaik.
5. Evaluasi merupakan Model data yang didapatkan dari alur datamining wajib dipaparkan menggunakan metode yang jelas bagi semua orang yang terlibat. tahap rendering dari proses KDD mencakup langkah ini. Pada tahap ini, kita menentukan apakah formula atau informasi yang didapatkan berlawanan pada kebenaran atau hipotesis yang telah diketahui.



Gambar 2. 1 Proses Data Mining Dalam Penemuan Pengetahuan Dalam Database

2.1.2 Data Mining

Menemukan pola, keteraturan, atau korelasi dalam volume data yang sangat besar adalah tujuan dari data mining, yang juga dikenal sebagai penemuan pengetahuan dalam basis data (KDD). Pada intinya, penemuan basis data (KDD) adalah proses penggalian informasi yang bernilai dari sekumpulan data. Penemuan basis data (KDD), menurut A. Berstein dkk., adalah produk akhir dari proses penelitian yang terdiri dari berbagai teknik algoritmik untuk menangani data, membuat model data, dan memodifikasi model. Data mining, seperti yang didefinisikan oleh Turban menjelaskan prosedur untuk menemukan informasi dalam database dalam bukunya *Decision Support Systems and Intelligent Systems*. Penambangan data menggunakan matematika, kecerdasan buatan, pembelajaran mesin, dan metode statistik untuk mengekstrak dan menemukan informasi yang relevan dari basis data yang cukup besar (Elisa 2018)

Data mining ialah alur yang terjadi secara terus menerus menganalisis serta menyerap data (informasi) memakai dua pola pembelajaran mesin. Definisi yang luas dari sebuah konsep dikembangkan melalui pengamatan contoh-contoh spesifik dari ide yang diteliti, dan proses ini dikenal sebagai pembelajaran induktif. KDD adalah singkatan dari *Knowledge Discovery in Databases*, yang merupakan penggalian data dengan menggunakan pendekatan ilmiah. Penggalian data adalah fase dalam proses KDD dalam situasi ini.

Satu dari sekian metode serta kegunaan *data mining* ialah:

1. evaluasi [prediktif]

2. pengelompokan [deskriptif]
3. Penemuan Aturan Asosiasi [deskriptif]
4. Regresi [prediktif]
5. Deteksi anomali (audio) Penambangan data dibagi menjadi beberapa bagian karena ini adalah proses berantai. Pengguna dapat berpartisipasi dalam fase-fase ini secara langsung atau tidak langsung melalui mediator KDD.

Fase pada *data mining* dikelompokkan kedalam enam tahapan:

1. Pembersihan Data yang menjadi fokus KDD harus dibersihkan terlebih dahulu sebelum penggalian data dapat dilakukan. Data duplikat harus dihilangkan, data yang terus berubah-ubah harus di check serta kekeliruan ketik pada data harus diperbaiki sebagai bagian dari proses pembersihan. Selain itu, proses "memperkaya" data saat ini-yaitu menambahkan data tambahan yang signifikan serta diperlukan untuk KDD, seperti data atau informasi dari luar-dilakukan.
2. Integrasi data Integrasi data adalah proses penggabungan informasi dari banyak database ke dalam satu database baru. Penggalian data sering kali membutuhkan informasi dari banyak database atau file teks sebagai tambahan dari satu database. Untuk kualitas yang secara unik mengidentifikasi entitas, seperti atribut Nama, Jenis Produk, Nomor Pelanggan, dan lainnya, integrasi data dilakukan. Integrasi data harus dilakukan dengan hati-hati karena kesalahan dapat menghasilkan hasil yang ambigu atau bahkan salah mengarahkan operasi selanjutnya. Sebagai

contoh, jika ternyata integrasi data berbasis jenis produk mencampur item dari banyak kategori, maka akan muncul hubungan antar item yang tidak nyata.

3. Pilihan data Hanya data yang berguna untuk analisis yang dicari dari database karena tidak semua data sering digunakan. Misalnya, hanya ID pelanggan yang harus dimasukkan dalam analisis keranjang belanja jika kemampuan berbelanja seseorang sedang diselidiki.
4. Modifikasi data Dalam data mining, data diubah atau dikonversi ke dalam format yang dapat diproses. format data tertentu harus ada sebelum menggunakan teknik data mining tertentu. Beberapa teknik umum, seperti analisis asosiasi, misalnya, hanya dapat menerima data input kategorikal. Untuk memisahkan data ke dalam beberapa periode waktu, angka numerik yang berurutan harus digunakan. Transformasi data adalah nama lain dari prosedur ini.
5. Ketika menerapkan teknik untuk mengungkap informasi yang tersembunyi dan tak ternilai dalam data, proses penambangan adalah langkah yang paling penting.
6. Evaluasi pola, yang mengidentifikasi pola-pola menarik dalam basis data yang ditemukan. Untuk menentukan apakah hipotesis yang sudah ada sebelumnya didukung, hasil dari teknik data mining diperiksa pada tahap ini dengan menggunakan pola umum dan model prediksi.
7. Information display (penyajian informasi), merupakan penyajian informasi dan penggambaran bagaimana pengguna menerima informasi yang telah

diterima. Perumusan penilaian atau tindakan berdasarkan hasil analisis yang diperoleh adalah tahap terakhir dari proses ini. Penerima temuan data mining terkadang tidak menyadari data mining itu sendiri. Karena semua orang dapat memahami informasi, maka sangat penting untuk menampilkan hasil data mining dengan cara ini. Model yang mudah dipahami ini akan menjadi langkah penting dalam data mining. Selain itu, visualisasi dapat berkontribusi pada komunikasi yang lebih efisien dan persuasif dari hasil data mining (Eska, 2016).

2.1.3 Klasifikasi

Salah satu teknik data mining yang paling populer yang digunakan dalam berbagai aplikasi adalah klasifikasi, yang merupakan pendekatan klasifikasi. Metode kategorisasi membutuhkan dua jenis masalah yang berbeda: data pelatihan dan data uji. Metrik partisi terstruktur untuk mengklasifikasikan data lain, atau data uji, dievaluasi oleh algoritma data mining dengan menggunakan data pelatihan. Ada dua data artikel teks yang digunakan, dibagi menjadi set pelatihan dan set pengujian (Ardi Ramdani., 2022).

Sserta sesuai (Setio, 2020) Proses penggalian data yang disebut klasifikasi membagi data ke dalam pengelompokan atau kategori yang telah ditentukan sebelumnya. Metode pembelajaran terawasi yang disebut klasifikasi menggunakan data pelatihan berlabel untuk membuat aturan yang mengkategorikan data uji ke dalam kelompok atau kelas tertentu. Pohon keputusan, pengklasifikasi berbasis aturan, jaringan saraf, mesin

pendukung, dan pengklasifikasi Bayesian naif adalah beberapa metode klasifikasi yang digunakan. Setiap metode menggunakan algoritma pembelajaran untuk menemukan contoh yang menawarkan korelasi terbaik antara properti data input dan ID kelas.

2.1.4 Pohon keputusan (*Decision tree*)

Berdasarkan (Tangkelayuk 2022) Algoritma pengambilan keputusan yang dikenal sebagai pohon keputusan cukup populer. Metode yang cocok untuk kategorisasi atau prediksi adalah pohon keputusan. Metode yang cocok untuk kategorisasi atau prediksi adalah pohon keputusan. Simpul akar, simpul internal, dan simpul terminal adalah tiga bagian utama dari sebuah pohon. Dari simpul akar hingga simpul daun (ujung), pohon keputusan terdiri dari simpul-simpul keputusan yang dihubungkan oleh cabang-cabang. Simpul akar, simpul internal, dan simpul terminal adalah tiga komponen utama dari paradigma pohon keputusan.

Pohon keputusan adalah diagram alir dengan struktur seperti pohon, di mana setiap simpul inti merepresentasikan properti yang dapat diuji, setiap cabang merupakan hasil pengujian, dan simpul daun merepresentasikan kelas atau distribusi kelas tertentu. Pohon keputusan digunakan untuk mengantisipasi data yang belum memiliki kelas data yang jelas dengan cara mengidentifikasi sekelompok pola atau fungsi yang mendeskripsikan dan membedakan kelas data yang satu dengan yang lainnya (Pribadi, Poningsih, Tambunan 2020).

2.1.5 Algoritma C4.5

Sebuah teknik untuk membuat pohon keputusan dalam proses pengambilan keputusan adalah algoritma C4.5. Metode C4.5 merupakan pengembangan dari algoritma pohon keputusan induktif ID3 (Iterative Dichotomizer 3). Pencipta ID3 adalah J. Ross Quinlan. Sampel pelatihan, label pelatihan, dan karakteristik adalah input yang diperlukan untuk prosedur algoritma ID3. Algoritma ID3 ditingkatkan dan disempurnakan menjadi metode C4.5. Metode C4.5 menggabungkan sejumlah inovasi, seperti kemampuan untuk mengelola nilai yang hilang, data yang persisten, dan prosedur pemotongan. Setiap objek yang diurutkan dalam pohon harus diverifikasi nilai entropinya sebelum pohon keputusan dapat dibangun. Entropi adalah pengukuran dalam teori informasi yang digunakan untuk menilai tingkat ketidakmurnian dan homogenitas kumpulan data. Nilai information gain (IG) dari setiap karakteristik ditentukan dari nilai entropi. Entropy (S) adalah ukuran jumlah bit yang diperkirakan diperlukan untuk memisahkan kelas (+ atau -) dari beberapa data acak dalam ruang sampel S. Entropy sebanding dengan jumlah bit yang diperlukan untuk menentukan kelas. Semakin banyak entropi yang digunakan untuk mengekstrak lapisan, semakin rendah nilai entropinya. Keunikan S dievaluasi dengan menggunakan entropi. Seseorang menggunakan atribut dengan nilai gain tertinggi dari atribut yang dapat diakses untuk memilih atribut asli. Gunakan rumus yang ditunjukkan pada persamaan berikut untuk menghitung gain.

$$Gain(A) = Entropi(S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi(S_i)$$

Rumus 2. 1 Perhitungn 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

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 2 berikut.

$$\text{Entropi } (S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

Rumus 2. 2 Rumus Entropy

Dimana :

S : himpunan kasus

A : fitur

N : jumlah partisi S

p_i : proporsi dari S_i terhadap S

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap-tiap nilai.
- c. Bagi kasus dalam cabang.

d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.(Elisa 2018)

Sesuai (Mardi 2017) Algoritma C4.5 mengacu pada jenis metode yang digunakan untuk membuat distribusi terstruktur berbasis Decision Tree yang bersifat prediktif. langkah-langkah berikut: Siapkan informasi pelatihan yang dikumpulkan ke dalam kategori kepemilikan dari catatan historis di masa lalu.

1. Pilih akar dari atribut Rute setelah menghitung nilai penguatan untuk setiap atribut; rute pertama akan memiliki nilai penguatan tertinggi. Dengan menggunakan rumus berikut ini, tentukan nilai entropi terlebih dahulu sebelum menghitung nilai gain properti.

Rumus menghitung nilai entropy

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

Rumus 2. 3 Rumus Entropy

Keterangan :

S = himpunan

A = kelengkapan yang dipakai

n = jumlah partisi pada bagian S

p_i = proporsi dari S_i ke S

- a. Kemudian hitung nilai gain menggunakan persamaan

$$Gain(A) = Entropi(S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi(S_i)$$

Rumus 2. 4 Perhitungan Gain

Keterangan :

S = himpunan

A = keterangan

n = jumlah partisi di keterangan atau kelengkapan bagian A

$|S_i|$ = jumlah yang ada pada partisi i

$|S|$ = jumlah yang ada pada S

- b. Kembali lagi ke cara 2 sampai semua *record* terpartisi.
- c. Jika semua catatan memiliki kelas N yang sama, atribut catatan tidak dipartisi ulang, dan catatan cabang tidak kosong, partisi pohon keputusan berakhir. (Mardi 2017)

2.1.6 WEKA (Waikato Environment for Knowledge Analysis)

Weka Universitas Waikato di Selandia Baru menciptakan koleksi WEKA yang berisi alat pembelajaran mesin yang berguna untuk tujuan pendidikan, penelitian, dan aplikasi lainnya. WEKA mampu menyelesaikan masalah data mining yang sebenarnya. dalam kategorisasi berdasarkan penggunaan pembelajaran mesin, khususnya. Semua sistem dapat menjalankan program ini karena dibuat dengan menggunakan teknik berorientasi objek dalam bahasa hirarki kelas Java (Saputra, 2023)

Kumpulan alat pembelajaran mesin yang berguna disebut WEKA. Waikato Environment for Knowledge Analysis, atau WEKA, adalah sebuah platform penelitian, pendidikan, dan perangkat lunak yang dikembangkan oleh Universitas Waikato di Selandia Baru. WEKA mampu menyelesaikan masalah penggalian data praktis, terutama kategorisasi yang mendasari teknik pembelajaran mesin. Program ini dapat digunakan di hampir semua platform dan dikembangkan dalam kelas-kelas Java hirarkis dengan fungsi-fungsi berorientasi objek. Pada berbagai tingkatan, WEKA mudah digunakan dan dipraktikkan. Metode pembelajaran modern diimplementasikan dan dapat digunakan untuk menganalisis kumpulan data langsung dari baris perintah. Pra-pemrosesan data, klasifikasi, regresi, pengelompokan, aturan asosiasi, dan alat visualisasi semuanya disertakan dalam WEKA. Tanpa mengembangkan kode peristiwa apa pun, pengguna dapat melakukan prapemrosesan data, mengirimkannya ke skema pembelajaran, dan memeriksa pengklasifikasi yang dihasilkan dan kinerjanya.

Pohon keputusan adalah salah satu pendekatan data mining yang paling menonjol yang dapat digunakan dalam WEKA. Salah satu teknik kategorisasi yang lebih menarik adalah pohon keputusan, yang memerlukan pembangunan pohon keputusan dengan node keputusan yang dihubungkan oleh cabang-cabang dari node akar ke node daun (akhir). Kualitas akan diuji pada simpul-simpul keputusan, dan setiap hasil akan menghasilkan sebuah cabang. Setiap cabang akan diarahkan ke node lain atau node pengambil keputusan. Pohon keputusan adalah salah satu teknik klasifikasi dari Maimon dan Rokach yang dinyatakan sebagai partisi rekursif (Singarimbun Hasibuan 2023).

2.2 Teori Khusus

Teori ini membahas istilah serta pengertian secara khusus berkaitan dengan pelaksanaan penelitian.

2.2.1 Kepuasan Pelanggan

Ketika membandingkan kinerja atau hasil yang diamati dengan harapan, kepuasan mengacu pada sejauh mana kualitas barang dan jasa yang dirasakan memenuhi harapan konsumen atau pembeli. Jika aktivitas menyatakan bahwa kesesuaian dan kemampuan untuk menggunakan karakteristik umum dari produk dan jasa yang diberikan untuk memenuhi kebutuhan dan harapan konsumen dengan menggunakan karakteristik atau faktor-faktor seperti: penampilan langsung, perhatian langsung karyawan kepada konsumen, daya tanggap, keandalan, dan jaminan. Pengalaman masa lalu, komentar keluarga, janji yang dibuat oleh pemasar, dan data yang diberikan oleh saingan, semuanya dapat mempengaruhi harapan klien (Riyani, 2021)

Pelanggan yang senang dengan sebuah bisnis akan sangat loyal dan tidak akan berpindah, bahkan merekomendasikannya kepada teman dan keluarga yang dapat menjadi pelanggan. Antisipasi untuk dapat mempercayai perkataan seseorang adalah apa yang dikenal sebagai kepercayaan (Elisa 2018)

2.2.2 Thrift

Industri penghematan adalah salah satu industri yang paling populer di masyarakat, terutama di kalangan anak muda. Istilah "thrift" sendiri berasal dari kata "thrive", yang berarti "berkembang atau maju", tetapi kata "thrifty" sendiri

berarti "menggunakan uang atau sumber daya lainnya dengan cara yang tepat dan efisien." Penghematan juga dapat dipahami sebagai tindakan membeli barang-barang lama, tetapi praktik ini juga mengacu pada kenikmatan pribadi seseorang dalam memperoleh produk berkualitas dengan anggaran yang lebih ketat atau dengan biaya yang lebih murah. Dari pakaian, topi, sepatu, dan celana hingga pakaian asing. Namun, ia juga menawarkan barang-barang dari perusahaan ternama lainnya, seperti Dickies, Nike, dan Adidas. Pemerintah telah melarang pembelian dan penjualan pakaian bekas impor karena beberapa alasan, termasuk kepedulian pemerintah terhadap citra yang sehat, meskipun baru-baru ini telah menjadi berita utama. Namun, saat ini pakaian bekas semakin disukai, terutama di kalangan anak muda. Semakin populernya industri perdagangan pakaian bekas sebagian disebabkan oleh kebutuhan masyarakat yang semakin mendesak. Ketika ada keuntungan yang dapat dirasakan oleh masyarakat terkait, industri pakaian bekas semakin berkembang karena adanya permintaan yang kuat dari pedagang dan pelanggan. Hal ini dikarenakan pakaian bekas dapat diolah kembali sehingga selalu memiliki nilai jual yang baik (richard oliver , 2021)

2.3 Penelitian Terdahulu

Peneliti juga menemukan beberapa hasil penelitian terdahulu yang berkaitan dengan menganalisa menggunakan algoritma C4.5 yang dirangkum yaitu :

1. Menurut penelitian Implementasi data mining menggunakan algoritma C.45 untuk memprediksi kepuasan pelanggan oleh (Nugraha.,2020) Dengan menggunakan aplikasi Rapid Minner, data pelanggan Gojek digunakan untuk membuat pohon keputusan untuk menerapkan algoritma

C4.5 pada data kepuasan pelanggan Gojek berdasarkan hasil penelitian. Hasil dari tiga kali penerapan algoritma C4.5 oleh penulis pada data kepuasan pelanggan Gojek menunjukkan bahwa pengujian 1, 2, dan 3 sudah baik dan berhasil karena nilai yang dihasilkan memiliki presisi yang tinggi dan meningkat, yang menandakan bahwa algoritma tersebut sudah akurat. Algoritma kepuasan pelanggan C4.5 dari Gojek dapat digunakan secara efektif.

2. Menurut penelitian yang berjudul "Analisis Kepuasan Masyarakat Terhadap Pelayanan Pengadilan Agama Pematangsiantar," Dengan menggunakan algoritma C.45, pendekatan ini membuat kategorisasi menjadi lebih sederhana, mudah, dan dengan akurasi yang tinggi. Penelitian ini dapat memberikan informasi bagi Pengadilan Agama Pematangsiantar dalam meninjau kebijakan dan pelayanan. Penelitian ini dilakukan oleh (Pribadi., 2020). Pengadilan Agama Pematangsiantar akan mendapatkan manfaat dari penelitian ini dengan mendapatkan pelayanan yang terbaik.
3. Penelitian "Indeks Kepuasan Pelanggan Informa Menggunakan Algoritma C.45 Amanda" oleh (Febriyani, Prayoga Nurdiawan, 2021) menemukan bahwa dengan menggunakan algoritma yang sama dengan C4.5 terbukti dapat mengukur kepuasan pelayanan, yang artinya pelayanan merupakan kunci keberhasilan karena jika perusahaan tidak memberikannya maka pelanggan akan meninggalkan perusahaan tersebut.

4. Informasi untuk pemeriksaan kebijakan dan pelayanan Pengadilan Agama Pematangsiantar dari penelitian Prediksi Tingkat Kepuasan Pelayanan Online Menggunakan Metode Algoritma C.45 yang diselesaikan olehn. Pada tahun 2020, penelitian ini telah selesai (Pribadi). Penelitian ini akan memberikan dukungan terbaik bagi Pengadilan Agama Pematangsiantar (Indah Lestari Defit, 2021) Di Dinas Kependudukan dan Catatan Sipil Kota Padang, informasi dikumpulkan untuk menilai kepuasan. Pendekatan penelitian menggunakan algoritma C4.5, yang merupakan metode prediksi.
5. Dari penelitian Elisa 2018 "Prediksi Laba pada Perusahaan dengan Klasifikasi Algoritma C4.5" Menurut data yang dikumpulkan, algoritma C4.5 menggunakan informasi dari anggaran konsultan perencana untuk meramalkan profitabilitas bisnis. Kita tahu bahwa informasi adalah sesuatu yang sangat penting dan berharga karena dengan menguasai pengetahuan kita dapat dengan mudah mencapai tujuan yang diinginkan, yang mendorong individu untuk memperebutkan informasi. Data mining adalah proses untuk mengekstraksi informasi baru dari sebuah tumpukan atau gudang data. Informasi diambil dari penelitian Penerapan Data Mining Algoritma C4.5 untuk Mengetahui Tingkat Kepuasan Pelanggan pada Hotel Grand Zuri Dumai yang dilakukan oleh (Desyanti 2018) Algoritma C45 digunakan untuk mengumpulkan data untuk menilai seberapa puas pelanggan terhadap layanan hotel Grand Zuri Dumai. Temuan penelitian ini adalah sebagai berikut: jika layanan sangat puas, keramahan sangat

ramah, kenyamanan menyenangkan, fasilitas memuaskan, dan pelanggan sangat puas.

6. Menurut penelitian "Penerapan Algoritma C.45 untuk Menentukan Tingkat Kepuasan Pelanggan Kartu Prabayar Telkomsel" yang dilakukan oleh (Sikumbang, 2022) Data yang dikumpulkan dengan menggunakan pendekatan algoritma C.45 untuk mengukur tingkat kepuasan pelanggan dalam menggunakan kartu prabayar Telkomsel.
7. Kesimpulan yang diambil dari penelitian "Data Mining Menggunakan Algoritma C4.5 untuk Memprediksi Kepuasan Mahasiswa terhadap Kinerja Dosen di Kota Batam" (Yulia anggia, 2019) Data yang dihasilkan menunjukkan bahwa algoritma C4.5 dapat digunakan untuk penggalan data untuk mengkategorikan data dan mencari tren untuk menghasilkan data yang mendalam untuk proses pengambilan keputusan. Selain itu, pohon keputusan dari aplikasi WEKA dan komputasi manusia dari output algoritma C4.5 menghasilkan pohon keputusan dengan tingkat akurasi 94,12%.
8. Menurut penelitian berjudul "Penerapan Algoritma C4.5 untuk Memprediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang" yang dilakukan oleh (Febriarini & Astuti, 2019) Ketika para ilmuwan ingin mengkategorikan atau mengelompokkan data dalam jumlah yang sangat besar, metode tipe C4.5 dapat memberikan hasil yang tepat.
9. Dari penelitian "Penerapan *Data Mining* Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5" yang dilakukan oleh

(Takalapeta 2018) Data yang diperoleh menunjukkan bahwa algoritme tipe C4.5 dapat membantu bisnis untuk berkonsentrasi dalam meningkatkan layanan pelanggan dan menawarkan wawasan yang signifikan dalam menetapkan kebijakan restoran.

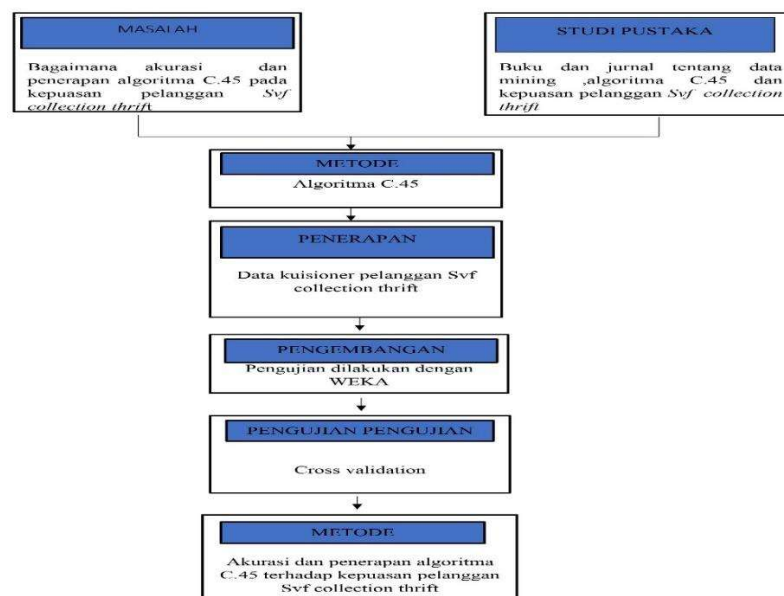
10. Dari penelitian "Algoritma C4.5 Untuk Mengetahui Kepuasan Jamaah Masjid Darul Hijrah Terhadap Program Dan Fasilitas DKM" yang dilakukan oleh (Warsino 2018) Informasi yang dikumpulkan melalui metode penggalian data telah menghasilkan temuan yang menjelaskan perilaku dan berhubungan dengan pendapat jamaah masjid tentang program, layanan, dan kualitas layanan. keunggulan layanan. Fasilitas dan sifat persekutuan setiap jemaah di masjid.
11. Dari Penelitian dengan judul "Klasifikasi *Data Mining* Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang" yang dilakukan (Faisal 2019) mendemonstrasikan penggunaan algoritme C4.5 dalam penggalian data yang memudahkan para ilmuwan untuk mendapatkan pengetahuan tentang kebahagiaan konsumen serta menghasilkan tingkat presisi yang tinggi, yaitu lebih dari 95%.
12. berdasarkan studi dari "Analisis Kepuasan Konsumen Terhadap Pelayanan Bengkel Menggunakan Metode Algoritma C4.5" (Alawiah,2021) mengklaim bahwa Bengkel Keluarga Zul Jaya Pematang Siantar menggunakan algoritma C4.5 untuk mengidentifikasi kepuasan klien dengan tingkat akurasi 94%.
13. Dari Penelitian dengan judul "Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5" yang dilakukan(Azwanti n Elisa 2020) dapat memberikan

aturan pohon keputusan yang mudah dipahami dan memiliki fungsi analisis kepuasan pelanggan, yang membuat algoritma C4.5 menjadi algoritma data mining yang berpotensi memberikan hasil yang baik.

14. Dari penelitian “Prediksi Tingkat Kepuasan Pelayanan Online Menggunakan Metode Algoritma C.45” yang dilakukan oleh (Indah Lestari & Defit, 2021) Informasi yang diperoleh menunjukkan bahwa algoritma C4.5 dapat secara akurat memprediksi suatu nilai atau menentukan tingkat kepuasan masyarakat pada tahun 2021.

2.4 Kerangka Pemikiran

Dibawah ini adalah kerangka pemikiran dari pengkajian dan juga penjelasannya.



Sumber : Peneliti 2023

Gambar 2. 2 Kerangka Pemikiran

Riwayat penjualan pakaian bekas di toko barang bekas Svf Collection selama kurang lebih satu tahun menjadi sumber data pembeli yang digunakan dalam survei, dan teknik algoritma C4.5 digunakan untuk memproses data. Program WEKA digunakan untuk menguji data, dan temuannya diperoleh sebagai ukuran kepuasan konsumen terhadap penjualan pakaian bekas di Svf Collection Thrift.

2.5 Hipotesis Penelitian

Algoritma C4.5 dapat digunakan untuk menganalisis kepuasan konsumen dalam penjualan pakaian bekas dengan cara berikut:

Hipotesis 1 : Prediksi bahwa Algoritma C4.5 dapat digunakan untuk menganalisis data kebahagiaan pelanggan dari toko-toko pakaian bekas untuk membantu menentukan elemen-elemen utama yang mempengaruhi tingkat kepuasan konsumen.

Hipotesis 2: Prediksi bahwa algoritma C4.5 dapat membantu mengkategorikan klien berdasarkan seberapa puas mereka dengan penjualan pakaian konsinyasi. dengan menggunakan data kepuasan pelanggan untuk menerapkan algoritma C4.5.