

BAB II

TINJAUAN PUSTAKA

2.1. Teori

Penulis akan membahas beberapa teori yang menjadi pedoman dasar dalam melaksanakan penelitian, yaitu dengan menjelaskan mengenai pengertian dan istilah yang berkaitan dengan judul penelitian.

2.1.1. Teori Umum

Penulis membahas beberapa teori umum yang berkaitan dengan pelaksanaan penelitian yang terdiri dari pengertian dan istilah secara umum.

2.1.1.1 KDD (*Knowledge Discovery In Database*)

KDD memiliki kepanjangan *Knowledge Discovery In Database* dan KDD adalah salah satu cara dalam mendapatkan dan pengetahuan dengan memanfaatkan data yang berasal dari *database* atau data hasil penyimpanan. Setelah menemukan pengetahuan tersebut, maka terakhir digunakan sebagai dasar pengetahuan dalam melakukan pengambilan keputusan. Demikian tahapan dalam KDD, yaitu:

1. *Data Selection*

Tahap ini dilakukan pada awal proses KDD yaitu dengan melakukan pengumpulan informasi dengan melalui proses seleksi data untuk digunakan sebagai sumber data yang pada akhirnya akan diolah pada tahap *data mining*.

2. *Pre-processing* atau *Cleaning*

Tahap ini dilakukan dengan tujuan menghilangkan beberapa duplikasi yang terdapat pada data, dengan memeriksa yang tidak sesuai dengan kebutuhan karena proses penambahan data harus sesuai dengan kebutuhan penelitian.

3. *Transformation*

Proses perubahan dan penyesuaian terhadap pola penyimpanan basis data.

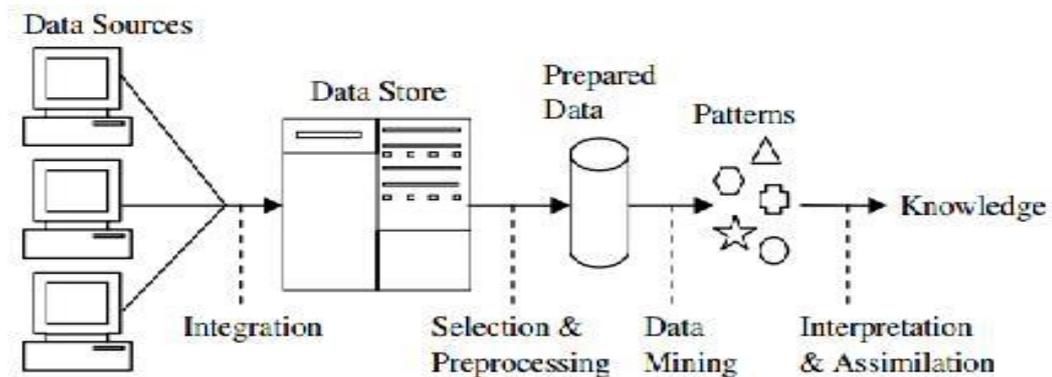
4. *Data Mining*

Proses yang menjadi cara dalam mencari dan membentuk pola aturan untuk menghasilkan informasi berupa keputusan berdasarkan tujuan penelitian.

5. *Interpretation*

Menampilkan hasil pola aturan dari *data mining* agar dapat dipahami, khususnya informasi yang bertentangan dengan hipotesis penelitian. (Mardi, 2016)

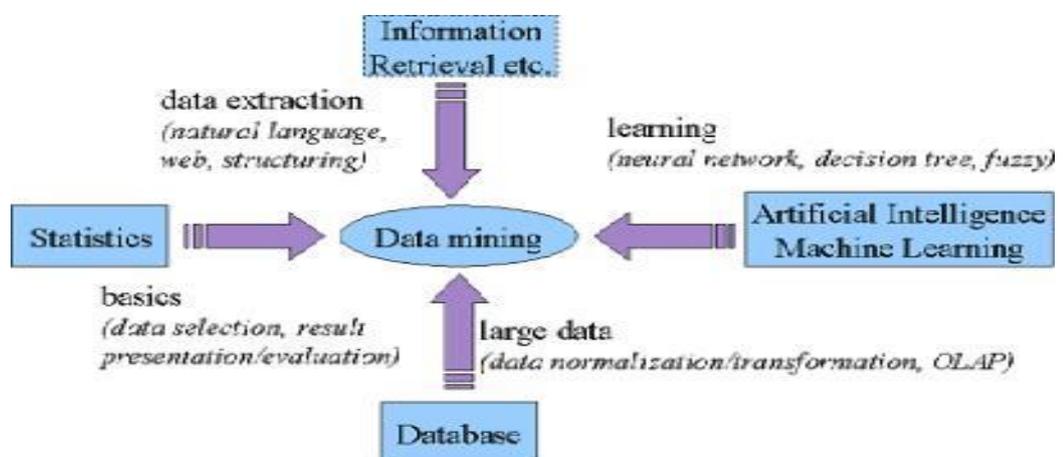
Sementara itu menurut (Mardi, 2016) proses KDD atau *Knowledge Discovery in Database*, yaitu:



Gambar 2. 1 Proses Dalam KDD atau *Knowledge Discovery in Database*
Sumber: (Mardi, 2016)

2.1.1.2 Data Mining

Data mining atau *machine learning* dan *data mining* bisa dimanfaatkan dan sering digunakan dalam melakukan berbagai bentuk kegiatan analisis untuk menjelaskan ilmu dan pengetahuan secara jelas dan mudah dipahami oleh pikiran yang dimiliki manusia (Hermawan, Sukma, & Halfis, 2019).



Gambar 2. 2 Bidang Ilmu Dari Data Mining
Sumber: (Mardi, 2016)

Dalam prosesnya, *data mining* memiliki teknik dalam pengolahan data tersendiri hingga terbentuk sebuah pola, kemudian pola tersebut difungsikan untuk dapat mengenali pola lain yang tidak tersimpan dalam lokasi penyimpanan data dan informasi yang sama. Dalam kegiatan atau proses melakukan prediksi, banyak peneliti melakukan kegiatan tersebut dengan memanfaatkan fungsi dari teknik *data mining*. Awal mula istilah *data mining* yaitu karena data dan informasi yang akan diproses sudah lama dan cara pengolahannya dilakukan seperti kegiatan menambang hingga mendapatkan hasil yang berharga dan bermanfaat (Hermawan et al., 2019).

2.1.1.3 Pengelompokkan *Data Mining*

Apabila dilakukan pembagian berdasarkan tahapan berdasarkan tugas dan proses yang dilakukan dalam *data mining* dapat dibagi menjadi beberapa tahapan dan bagian, yaitu:

1. *Description* atau Deskripsi

Dalam tahap ini, bermula ketika orang yang sedang melakukan kegiatan analisa berusaha untuk mengelompokkan jenis pola yang terdapat dalam sumber data dan informasi. Hal ini dapat digambarkan seperti pada saat seorang petugas yang bertanggung jawab dalam melakukan kegiatan penghitungan pengumpulan suara dalam kegiatan pemilihan tidak mengetahui atau mendapatkan keterangan yang jelas mengenai baik atau tidaknya kemampuan serta keahlian yang dimiliki oleh calon yang terdaftar

dalam pemilihan. Sehingga deskripsi yang dimaksud adalah upaya dalam memberikan beberapa penjelasan akan kemungkinan mengenai kemampuan dari calon yang terdaftar dalam pemilihan tersebut dengan suatu pola.

2. *Estimation* atau Estimasi

Selanjutnya, dalam tahapan atau proses ini kegiatan yang dilakukan dari estimasi memiliki cara kerja yang mirip dengan proses atau tahap dari klasifikasi, namun dapat berbeda apabila variabel dari target yang terdapat dalam proses atau tahapan dari estimasi berisi data yang bersifat numerik. Pola yang dibuat berasal dari rekaman data dengan berisi nilai atau numerik dari variabel yang terdapat dalam untuk dijadikan bahan dalam memberi nilai untuk proses dari prediksi. Kemudian dilanjutkan dengan pemberian nilai dari estimasi yang berasal dari nilai yang terdapat dalam variabel yang bersumber dari nilai dalam prediksi. Hal tersebut dapat digambarkan dengan nilai yang terdapat dalam IPK atau Indeks Prestasi Kumulatif mahasiswa yang sedang menempuh program pendidikan strata satu (S.1) dapat diketahui nilai estimasinya berdasarkan hasil nilai dari IPK atau Indeks Prestasi Mahasiswa pada saat proses pendidikan setiap semesternya.

3. *Prediction* atau Prediksi

Dalam proses atau tahapan dari prediksi juga hampir mirip dengan proses dan tahapan pada saat melakukan klasifikasi dan juga pada saat melakukan proses estimasi, namun tidak termasuk apabila prediksi dari nilai yang ada diperuntukkan untuk masa yang akan datang. Prediksi yang dimaksud khususnya berkaitan dalam kegiatan bisnis dan juga penelitian, yaitu:

- a. Pada saat melakukan prediksi untuk mengetahui kemungkinan terhadap harga dari beras selama beberapa bulan yang akan datang.
- b. Pada saat memberikan prediksi mengenai kemungkinan yang terjadi mengenai pengangguran selama beberapa tahun kedepan.
- c. Pada saat membuat prediksi dalam bentuk persentase yang membahas mengenai kemungkinan akan kecelakaan yang akan terjadi apabila pengendara mengemudi kendaraannya melawati batas kecepatan yang disarankan selama beberapa tahun kedepan.

4. *Classification* atau Klasifikasi

Dalam proses atau kegiatan saat klasifikasi, digunakan beberapa target dalam variabel berdasarkan kategori yang ada. Hal tersebut dapat digambarkan seperti pada saat melakukan pengelompokkan mengenai penghasilan yang dimiliki oleh pekerja dalam sebuah perusahaan, pengelompokkan tersebut dilakukan dengan menjadikan penghasilan pekerja memiliki tiga kategori, yaitu pekerja dari kategori penghasilan yang sangat tinggi, pekerja dari kategori tingkat penghasilan yang sedang atau standar sedangkan pekerja dari kategori tingkat penghasilan yang sangat rendah. Kemudian apabila klasifikasi dijelaskan dengan kaitannya dengan kegiatan penelitian dan dalam bisnis, yaitu:

- a. Pada saat melakukan proses klasifikasi terdapat transaksi kredit yang dilakukan dengan curang atau tidak.
- b. Pada saat melakukan prakiraan terhadap rencana kredit dilakukan oleh nasabah dengan tindakan yang tepat atau tidak.

- c. Pada saat dokter melakukan diagnosa terhadap penyakit atau kelainan yang dialami dan dimiliki pasien berdasarkan gejala yang ada.

5. *Clustering* atau Pengklusteran

Proses dari tahapan yang terdapat dalam pengklusteran adalah kegiatan atau proses melakukan pengelompokan hasil perekaman data yang kemudian dikeompokkan lagi berdasarkan data yang memiliki kesamaan atau kemiripan pola. Kluster yang dimaksud memiliki hasil perekaman data yang memiliki suatu kemiripan atau kesamaan pola dalam satu kelompok dan kelompok lainnya tidak memiliki kesamaan atau kemiripan dengan hasil perekaman yang ada. Dalam proses atau tahapan yang terdapat dalam pengklusteran berbeda dengan proses atau tahapan dari klasifikasi karena dalam pengklusteran tidak terdapat variabel dari target. Cara kerja yang dilakukan dalam tahapan dan proses dalam pengklusteran tidak melakukan proses yang terdapat pada saat melakukan estimasi, prediksi dan juga klasifikasi karena perhitungan dari algoritma yang digunakan dalam proses pengklusteran fokus melakukan pengelompokkan data yang ada berdasarkan kesamaan pola yang dimiliki oleh data dan untuk pemberian nilai yang maksimal berasal atau sesuai dengan tingkat kemiripan atau kesamaan data yang dimiliki dalam kelompok begitu juga berlaku dengan sebaliknya apabila tingkat kesamaan atau kemiripan data yang ada sangat rendah maka hasil dari nilai tersebut menjadi minimal. Hal tersebut dapat digambarkan, sebagai berikut:

- a. Pada saat melakukan pengelompokan terhadap jenis-jenis dari konsumen yang dijadikan target dari kegiatan pemasaran dari sebuah produk yang dilakukan suatu perusahaan.
- b. Pada saat melakukan kegiatan audit, dilakukan pemisahan berdasarkan transaksi yang diduga mencurigakan dan transaksi yang diduga aman.

6. *Association* atau Asosiasi

Asosiasi melakukan prosesnya untuk mendapatkan kelengkapan atau keterangan pendukung yang terdapat dalam sebuah data. Hal tersebut dapat digambarkan melalui bentuk kegiatan yang dilakukan, yaitu:

- a. Pada saat melakukan penelitian atau analisa mengenai jumlah pelanggan dari sebuah perusahaan yang memberikan respon secara positif terhadap pelayanan yang diberikan perusahaan.
- b. Untuk mengetahui keseluruhan jenis barang yang ada dipasar namun belum pernah dibeli atau dikonsumsi sebelumnya. (Mardi, 2016)

Penjelasan dalam bentuk gambar mengenai hasil pengelompokan beberapa proses atau tahapan berdasarkan tugas saat seseorang peneliti melakukan penelitian menggunakan teknik dengan *data mining*, sebagai berikut.



Gambar 2. 3 Pengelompokan Teknik Data Mining
Sumber: (Hermawan et al., 2019)

2.1.1.4 *Decision Tree*

Akibat dari berbagai jenis permasalahan dan tingkat kesulitan dalam permasalahan yang berbeda dan sering terjadi dalam kehidupan manusia pada akhirnya membuat manusia mulai melakukan berbagai upaya dalam mengembangkan sebuah pola berbentuk sistem agar dapat membantu mereka dalam menyelesaikan berbagai jenis dan bentuk permasalahan yang terjadi dan salah satu hasilnya adalah *decision tree* atau biasa dikenal dengan istilah pohon keputusan. *Decision tree* merupakan salah satu bentuk implementasi pola dengan menerapkan berbagai tahapan dan proses yang terdapat dalam kegiatan klasifikasi dan juga kegiatan prediksi. *Decision tree* menghasilkan alternatif dalam memecahkan masalah dengan hasil yang kuat dan berguna sehingga *decision tree* terkenal dan banyak digunakan oleh peneliti saat melakukan penelitian dengan teknik *data*

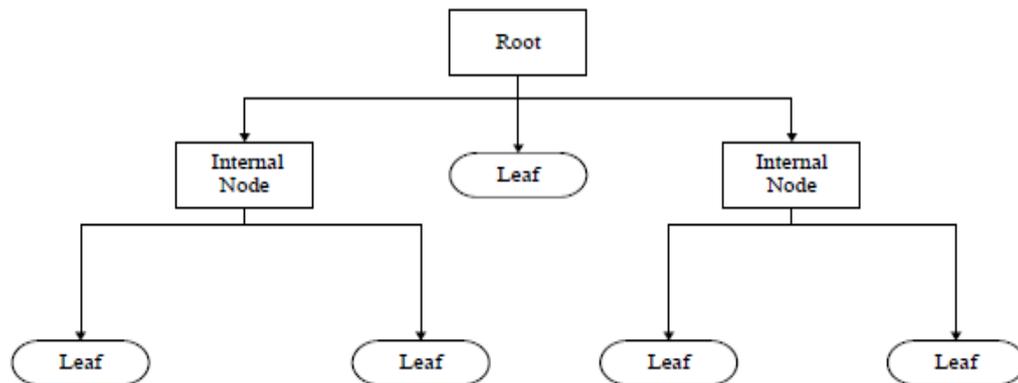
mining. *Decision tree* sangat bermanfaat apabila digunakan untuk keperluan dalam mengubah data yang ada menjadi sebuah pola berbentuk aturan sehingga dapat dipahami dengan mudah oleh manusia. Dalam tahapannya, *decision tree* melakukan proses yang dimulai dari keseluruhan bagian-bagian yang terdapat dalam data yang berisi berbagai kondisi dan keterangan yang lengkap dan kemudian diakhiri dengan sebuah keputusan. Arsitektur dari *decision tree* memang dibuat mirip dengan struktur yang terdapat pada sebuah pohon dan penjelasannya, yaitu:

1. Simpul bagian akar
Simpul dari bagian akar terletak pada bagian paling atas dan pertama dari struktur yang terdapat pada *decision tree*.
2. Simpul internal
Simpul dari bagian internal ini adalah percabangan lanjut dari simpul bagian akar. Dalam sebuah simpul ini apabila terdapat satu masukan maka dapat mengeluarkan keluaran dengan jumlah maksimal sebanyak dua.
3. Simpul daun
Simpul daun merupakan simpul yang berada pada ujung bagian dari pohon. Dalam simpul daun terdapat sebuah masukan namun tidak memiliki keluaran. Selanjutnya, demikian penjelasan dari kelebihan dan kekurangan yang dimiliki oleh *decision tree*, yaitu:

1. Kelebihan dari *decision tree*
 - a. Dalam penggunaan *decision tree* apabila terdapat cakupan data yang rumit maka dirubah menjadi lebih sederhana untuk kemudian dijadikan dasar pengambilan keputusan.

- b. Kelebihan dari penggunaan *decision tree* yaitu dapat menghapus cara perhitungan yang tidak perlu sehingga dalam melakukan proses dari pengujian hanya menggunakan kriteria sesuai dengan keperluan untuk pengujian saja.
 - c. Dalam proses pemilihan akan kelengkapan atau keterangan pada data, bila menggunakan *decision tree* maka hasil pemilihan terhadap kelengkapan atau keterangan data tersebut akan menjadi pembeda dalam mengelompokkan data berdasarkan kriteria lainnya.
 - d. *Decision tree* bertindak untuk menghindari permasalahan lain dengan memanfaatkan menggunakan jenis kelompok atau kategori dengan jumlah sedikit tanpa mengurangi kualitas hasil keputusan akhir.
2. Kekurangan dari *decision tree*
- a. Dalam menggunakan *decision tree* apabila jumlah dari hasil keputusan yang harus dihasilkan dan kriteria yang digunakan memiliki jumlah yang sangat banyak, maka waktu yang dibutuhkan dalam menghasilkan keputusan akan semakin bertambah serta dibutuhkan memori sistem pendukung yang tinggi.
 - b. Dalam *decision tree* terdapat akumulasi dari jumlah *error* yang terjadi dalam setiap bagian dari struktur yang terdapat pada *decision tree*.
 - c. Membuat struktur dari *decision tree* dengan hasil yang maksimal ternyata sulit untuk dilakukan.

- d. Apabila menggunakan *decision tree*, maka hasil atau kualitas dari keputusan yang dihasilkan sangat ditentukan dari model atau struktur *decision tree* yang dibuat. (Harryanto & Hansun, 2017)



Gambar 2. 4 Model Decision Tree
Sumber: (Takalapeta, 2018)

2.1.1.5 Algoritma C4.5

Algoritma dari C4.5 bermanfaat dalam melakukan proses klasifikasi data yang menggunakan angka atau yang menggunakan kategori. Setelah tahapan dari proses klasifikasi dilakukan maka kemudian hasilnya adalah beberapa aturan dari pola dan dapat dipergunakan untuk melakukan tahapan dari proses prediksi dengan nilai-nilai prediksi yang berasal dari perekaman data yang lama hingga terbaru. Jenis algoritma dari C4.5 berasal dari algoritma ID3, maksudnya adalah algoritma dari C4.5 adalah hasil dari perbaikan lebih lanjut yang dilakukan terhadap algoritma dari ID3 sehingga dalam algoritma C4.5 memiliki kemampuan dalam mengatasi permasalahan saat terjadi ketiadaan data dan mengatasi permasalahan saat terjadi data yang bersifat pecahan. Pada umumnya algoritma dari C4.5 digunakan dalam

proses pembangunan struktur dari *decision tree*, dan demikian penjelasan proses atau tahapannya yaitu:

1. Pertama adalah memilih kelengkapan atau keterangan data sebagai simpul akar dari *decision tree*.
2. Kemudian dilanjutkan dengan membuat struktur cabang dari *decision tree* yang sudah berisi nilai-nilai.
3. Pada tahap ketiga, dilakukan pembagian atau pengelompokkan dari permasalahan yang terdapat pada struktur cabang dari *decision tree*.
4. Dan tahap akhir adalah melakukan proses pengulangan yang ada pada tahapan pertama hingga ketiga sampai seluruh cabang pada struktur *decision tree* memiliki kelas dan kategori yang sama.

Dalam melakukan tahap awal dari *decision tree* yaitu pada saat melakukan pemilihan terhadap kelengkapan atau keterangan data dilakukan dengan cara mencari nilai dari gain dengan hasil yang tertinggi dari keluruhan nilai yang dimiliki oleh kelengkapan atau keterangan yang ada. Demikian rumus untuk melakukan pencarian nilai gain tertinggi, yaitu:

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

Rumus 2. 1 Perhitungan Gain

Keterangan simbolnya, yaitu:

S adalah himpunan

A adalah keterangan atau kelengkapan yang digunakan.

n adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian A

$|S_i|$ adalah jumlah dari kasus yang terdapat dalam partisi yang ke- i

$|S|$ adalah jumlah dari kasus yang terdapat dalam S

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

Rumus 2. 2 Perhitungan Entropy

Keterangan simbolnya, yaitu:

S adalah himpunan

A adalah keterangan atau kelengkapan yang digunakan.

n adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian S

p_i adalah proporsi dari S_i kepada S (Elisa, 2017).

2.1.2. Teori Khusus

Penulis membahas beberapa teori khusus yang berkaitan dengan pelaksanaan penelitian yang terdiri dari pengertian dan istilah secara khusus.

2.1.2.1 Kepuasan Pelanggan

Rasa puas pelanggan bersumber dari prakiraan atau harapan dengan sebuah rasa yakin yang dimiliki oleh pelanggan sesaat sebelum membeli sebuah produk atau menggunakan jasa. Biasanya harapan bersumber kualitas yang diberikan produk atau prakiraan terhadap pelayanan jasa yang akan diterimanya disaat setelah membeli produk atau menggunakan berbentuk jasa tersebut. Pada saat setelah membeli atau melakukan konsumsi terhadap sebuah produk atau jasa, apabila pelanggan tidak suka dengan kualitas dari produk atau juga pelayanan maka

pelanggan tidak puas. Dan juga berlaku sebaliknya, apabila kualitas yang dirasakan dari hasil produk atau pelayanan sesuai dan pelanggan merasa suka maka kemungkinan pelanggan akan puas.

Biasanya sering dijumpai kejadian, apabila seorang pelanggan merasa puas terhadap kualitas sebuah produk atau juga pelayanan jasa maka pelanggan tersebut akan saling berbagi cerita atau pengalaman dengan rekan atau saudara terdekatnya, sehingga apabila hal tersebut terjadi maka kemungkinan produsen penghasil produk dan juga jasa tersebut akan mengalami keuntungan karena kemungkinan besar rekan atau saudara dari pelanggan yang merasa puas tersebut akan melakukan pembelian ulang atau kedua kali sehingga pada akhirnya pelanggan dan juga produsen dari produk atau jasa tersebut akan sama-sama saling merasakan keuntungan (Shiddiq, Niswatin, & Farida, 2018).

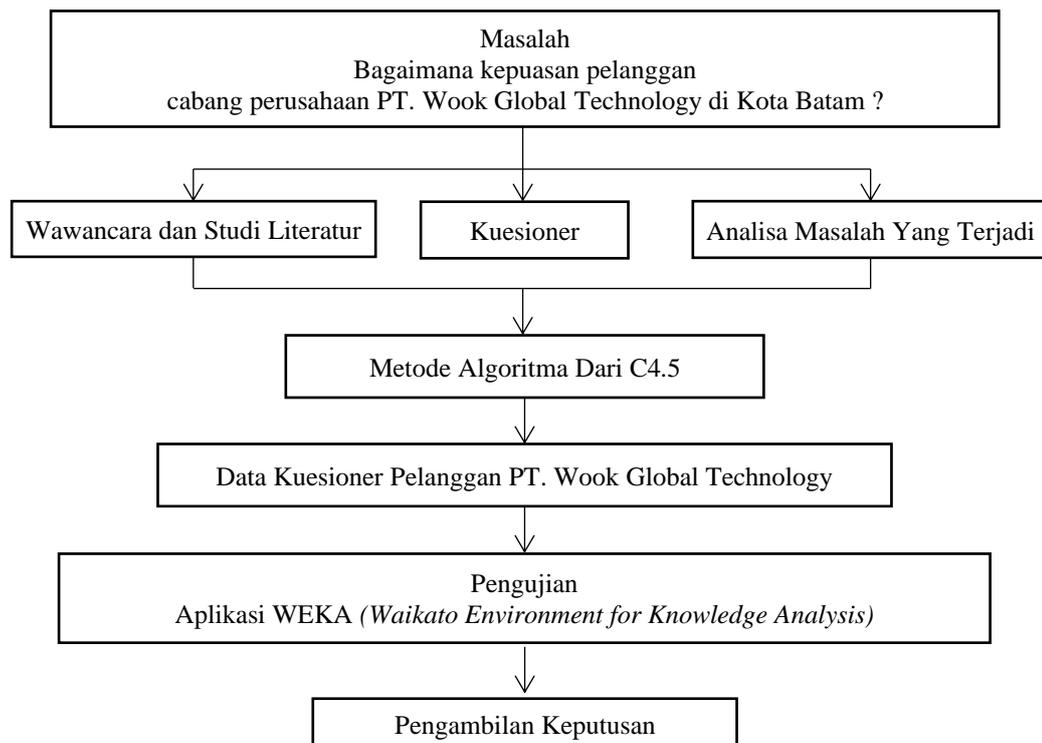
2.1.2.2 Aplikasi WEKA

Pada saat seorang peneliti ingin melakukan proses atau kegiatan perhitungan terhadap klasifikasi dari data dengan model atau jenis skala ordinal maka biasanya peneliti menggunakan aplikasi dari WEKA atau *Waikato Environment for Knowledge Analysis* (Effendy & Purbandini, 2018).

Aplikasi WEKA adalah sistem yang digunakan untuk melakukan pekerjaan dalam melakukan perbandingan terhadap algoritma dari mesin pembelajaran atau *machine learning* dalam penggunaan teknik *data mining* (Effendy & Purbandini, 2018).

2.2. Kerangka Pemikiran

Model kerangka dari pemikiran yang ada berdasarkan penelitian yang dibuat, yaitu sebagai berikut:



Gambar 2. 5 Kerangka Pemikiran
Sumber: (Penulis,2019)

2.3. Hipotesis Penelitian

Demikian penjelasan dari hipotesis yang menjadi atau dugaan penulis sementara terhadap penelitian, yaitu:

1. *Data mining* dengan algoritma dari C4.5 dapat menganalisis kepuasan pelanggan cabang perusahaan PT. Wook Global di Kota Batam.

2. Hasil penelitian dapat membantu pemecahan masalah dalam pengambilan keputusan mengenai kepuasan pelanggan terhadap PT. Wook Global Technology di Kota Batam dengan proses *data mining* dan perhitungan algoritma C4.5.

2.4. Penelitian Terdahulu

Informasi berupa referensi dan hasil penelitian dari peneliti lainnya namun tetap berkaitan dengan tema dan judul dari penelitian.

1. Dari penelitian “Penerapan *Data Mining* Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5” yang dilakukan oleh (Takalapeta, 2018) didapatkan informasi bahwa dengan penggunaan jenis algoritma dari C4.5 terbukti dapat membantu pihak dari restoran dalam fokus peningkatan pelayanan terhadap konsumen dan memberikan masukan informasi yang penting dalam menetapkan kebijakan.
2. Dari penelitian “Analisa Kepuasan Konsumen Menggunakan Klasifikasi *Decision Tree* Di Restoran Dapur Solo (Cabang Kediri)” oleh (Shiddiq et al., 2018) didapatkan informasi bahwa *decision tree* dengan algoritma dari C4.5 dalam mengolah data yang dimiliki restoran memberikan hasil berupa pola dalam bentuk aturan untuk mengetahui hal-hal yang dapat membuat konsumen puas dan tidak puas.
3. Dari penelitian “Algoritma C4.5 Untuk Mengetahui Kepuasan Jamaah Masjid Darul Hijrah Terhadap Program Dan Fasilitas DKM” yang dilakukan oleh

(Warsino, 2018) didapatkan informasi bahwa pada saat melakukan proses dari *data mining* dapat diketahui hasil yang menjelaskan mengenai perilaku dan juga mengenai persepsi dari jamaah yang berasal dari Masjid terhadap kualitas dari program, kualitas dari pelayanan, dan manfaat dari fasilitas serta bentuk hubungan dari setiap jamaan yang terdapat dalam Masjid.

4. Dari penelitian "*Data Mining Menggunakan Algoritma C4.5 Untuk Memprediksi Kepuasan Mahasiswa Terhadap Kinerja Dosen Di Kota Batam*" yang dilakukan oleh (Yulia & Putri, 2019) didapatkan informasi bahwa hasil prediksi mengenai kinerja Dosen dengan menggunakan dasar nilai tertinggi yang terdapat pada gain. Kemudian struktur *decision tree* yang dibuat dengan model perhitungan secara manual dengan menggunakan jenis algoritma dari C4.5 menghasilkan pola berbentuk aturan yang dapat menjelaskan proses yang terjadi dalam tahapan prediksi yang berhubungan dengan variabel dari penelitian yang dilakukan. Penggunaan atau cara *data mining* dengan menggunakan jenis algoritma C4.5 secara manual ternyata terbukti memiliki hasil berupa nilai akurasi yang sama dengan penggunaan aplikasi WEKA.
5. Dari penelitian "*Taiga: Performance Optimization of the C4.5 Decision Tree Construction Algorithm*" yang dilakukan oleh (Yang & Chen, 2016) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 yang terdapat pada *decision tree* dapat dimaksimalkan dengan melakukan penyelasaran data dengan kategori atau kriteria dalam melakukan *data mining* sehingga waktu yang dibutuhkan untuk mendapatkan keputusan akhir menjadi lebih singkat.

6. Dari penelitian "*Reservoir Inflow Forecasting Using ID3 and C4.5 Decision Tree Model*" yang dilakukan oleh (Charoenporn, 2017) didapatkan informasi bahwa jenis algoritma dari ID3 dan juga algoritma dari C4.5 merupakan salah satu cara dalam membuat gudang dalam melakukan pengujian dengan suatu model atau tahapan kerja. Maksudnya adalah jenis algoritma dari ID3 dan juga jenis algoritma dari C4.5 merupakan bentuk perhitungan untuk *decision tree*. Pada akhirnya dapat menghasilkan solusi atau keputusan akhir yang terbaik berdasarkan berbagai proses penyesuaian data.
7. Dari penelitian "*Selection of Appropriate Candidates for a Type Position Using C4.5 Decision tree*" yang dilakukan oleh (Rabcan, Vaclavkova, & Blasko, 2017) didapatkan informasi bahwa didapatkan hasil penelitian dengan tingkat akurasi yang tinggi dan meningkatkan hasil klasifikasi.
8. Dari penelitian "Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang *Bus Rapid Transit* (BRT) Trans Semarang yang dilakukan oleh (Febriarini & Astuti, 2019) didapatkan informasi bahwa penggunaan jenis Algoritma dari C4.5 sangat bermanfaat saat peneliti ingin melakukan proses klasifikasi atau pengelompokkan data dengan jumlah yang sangat banyak dan dapat mengasilkan nilai yang akurat .
9. Dari penelitian "Prediksi Kelayakan Masuk Jurusan IPA Siswa Sekolah Menengah Atas Menggunakan C4.5 STudi Kasus: SMA Tarakanita Gading Serpong) yang dilakukan oleh (Rahwali, Hansun, & Wiratama, 2017) didapatkan informasi bahwa hasil implementasi menggunakan *data mining* dan penggunaan algoritma dari C4.5 dalam kegiatan dan tujuan untuk

memprediksi kelayakan siswa dan siswi saat proses penjurusan dapat berhasil dilakukan dengan hasil yang sesuai dengan kebutuhan dan hasil yang akurat.

10. Dari penelitian “Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada PT. Capella Dinamik Nusantara Cabang Muka Kuning yang dilakukan oleh (Azwanti, 2018) didapatkan informasi bahwa jenis algoritma dari C4.5 bermanfaat dalam kegiatan atau proses melakukan klasifikasi kemudian juga dapat digunakan untuk tujuan prediksi penjualan dari motor. Pada tahap atau proses pengujian dengan aplikasi WEKA ternyata juga memberikan kemudahan bagi peneliti.