

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Teori Dasar**

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. Sistem Pendukung Keputusan (SPK)**

Menurut (Sri Lestari & Saefudin, 2015), Sistem pendukung keputusan merupakan bagian dari sistem informasi berbasis komputer, termasuk sistem berbasis pengetahuan atau manajemen pengetahuan, yang digunakan untuk mendukung pengambilan keputusan dalam suatu organisasi atau perusahaan. Pada dasarnya, SPK merupakan pengembangan lebih lanjut dari sistem informasi manajemen terkomputerisasi, yang dirancang sedemikian rupa sehingga dapat berinteraksi dengan pemakainya. Sistem pendukung keputusan terdiri dari tiga komponen utama yaitu subsistem pengolah data (database), subsistem pengolah model (modelbase) dan subsistem pengolah dialog.

### 2.1.1.2.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.

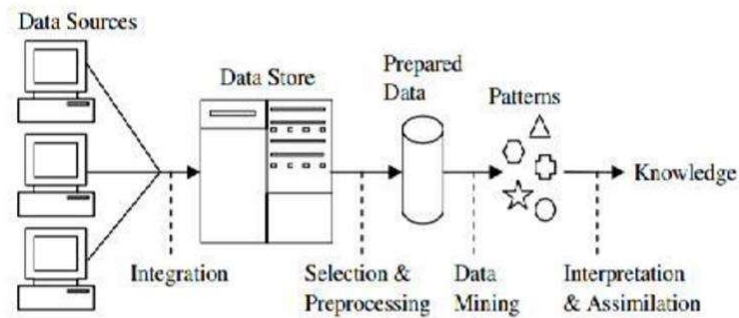
4. *Data Mining*

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:

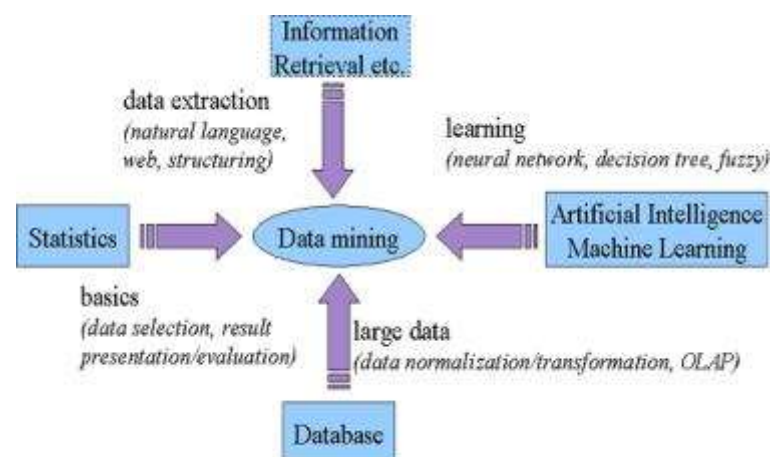


**Gambar 2.1** Proses Dalam KDD atau *Knowledge Discovery in Database*

**Sumber:** (Mardi, 2016)

### 2.1.1.3. 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 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).

#### **2.1.1.4. Pengelompokan *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 pengelompokan mengenai penghasilan yang dimiliki oleh pekerja dalam sebuah perusahaan, pengelompokan 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:

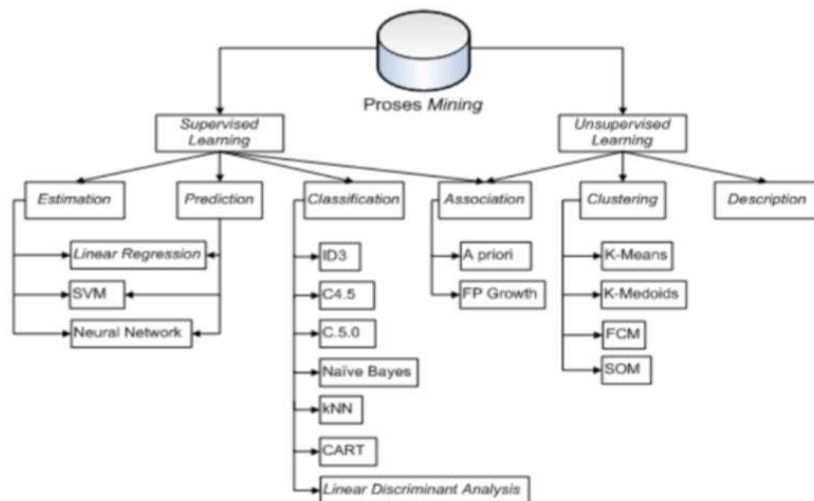
- a. Pada saat melakukan pengelompokkan terhadap jenis-jenis dari konsumen yang dijadikan target dari kegiatan pemasaran dari sebuah produk yang dilakukan suatu perusahaan.
- b. Pada saat melakukan kegiatan audit, dilakukan pemisahan berdasarkan transaksi yang diduga mencurigakan dan transaksi yang diduga 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, Sukma, & Halfis, 2019)

#### 2.1.1.5. *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:

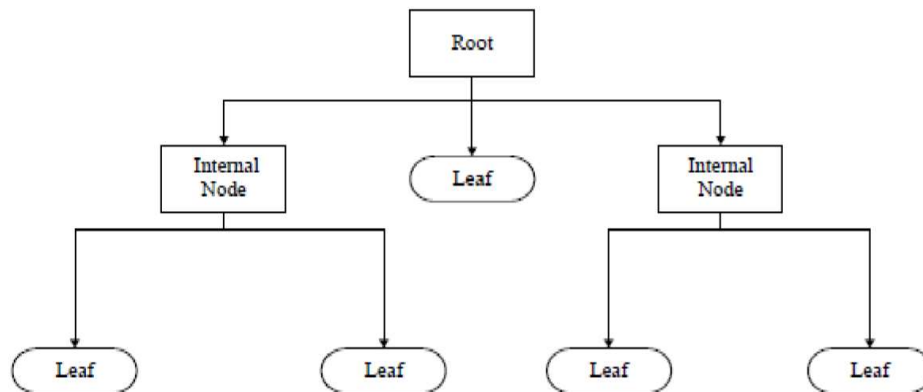
1. Kelebihan dari *decision tree*

- 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 penggunaan 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)

#### 2.1.1.6. 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:

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

#### 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  $S$

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

### 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. Kinerja Karyawan

Kinerja karyawan adalah kuantitas dan / atau kualitas pekerjaan yang dihasilkan oleh karyawan PT Indoland Batam, seperti standar kerja, dalam kurun waktu tertentu, berdasarkan peraturan yang ditetapkan atau berlaku, prosedur operasi, standar, dan fungsi di dalam perusahaan. Indikator yang digunakan untuk mengukur kinerja karyawan (Fransisca Andreani & Tanto Wijaya, 2015) adalah:

1. Jumlah pekerjaan yang di hasilkan
2. Kualitas pekerjaan yang di hasilkan
3. Ketepatan waktu menyelesaikan pekerjaan
4. Kehadiran pada hari dan jam kerja
5. Kemampuan kerja sama

#### **2.1.2.2.Reward**

*Reward* merupakan salah satu cara untuk memotivasi seseorang agar berbuat baik dan meningkatkan prestasinya. *Reward* adalah imbalan, ganjaran, hadiah, atau penghargaan yang dirancang untuk membuat seseorang lebih aktif bekerja untuk meningkatkan atau meningkatkan kinerja yang dicapai. *Reward* dapat diartikan sebagai bentuk penghargaan atas prestasi tertentu yang diberikan oleh individu atau organisasi, biasanya dalam bentuk materi atau verbal (Fitri, Ludigdo & Djamhuri, 2013). *Reward* yang diberikan kepada karyawan akan berdampak signifikan terhadap kinerja.

#### **2.1.2.3.Bonus**

Menurut (Agusli, Dzulhaq, & Khasanah, 2017), Bonus adalah pendapatan tambahan yang berfungsi sebagai pendorong semangat agar karyawan semakin bergairah dalam meningkatkan prestasi kerja dan loyalitas pada perusahaan. Adapun tujuan program bonus yang diberikan perusahaan, sebagai berikut:

1. Bagi Perusahaan

Tujuan pemberian bonus dimaksudkan untuk meningkatkan kinerja dengan cara mendorong karyawan agar bekerja disiplin dan memiliki semangat yang lebih tinggi. Diharapkan dengan pemberian bonus ini dapat mencapai tujuan

dalam menghasilkan kualitas kinerja karyawan yang lebih baik, meningkatkan efektivitas dan efisiensi dalam penggunaan faktor kinerja, serta mencegah terjadinya penurunan kinerja karyawan.

## 2. Bagi Karyawan

Dengan pemberian bonus dari perusahaan, diharapkan karyawan memperoleh banyak keuntungan. Keuntungan yang merupakan komponen non-upah ini, misalnya mendapatkan gaji yang lebih besar di akhir tahun, mendapat dorongan untuk mengembangkan dirinya, berusaha bekerja dengan sebaik-baiknya, dan loyalitas terhadap perusahaan.

### **2.1.2.4.Karyawan**

Karyawan merupakan faktor pendukung perusahaan atau instansi, karena dengan karyawan dengan standar kualifikasi perusahaan maka produktivitas perusahaan pasti akan terjaga dan ditingkatkan. Proses pemilihan karyawan yang unggul adalah proses kompleks yang membutuhkan pertimbangan yang cermat. Untuk memperoleh informasi yang cepat dan akurat tentang kinerja karyawan yang benar (sesuai dengan standar yang diharapkan) diperlukan proses otomatis dengan menggunakan teknologi. Oleh karena itu untuk memenuhi kebutuhan akan informasi diperlukan suatu sistem yang berbasis komputer (Afriany & Aisyah, 2019).

### **2.1.2.5.Insentif**

Insentif merupakan reward yang diberikan kepada pegawai oleh pimpinan organisasi berupa uang agar pegawai memiliki motivasi dan prestasi yang tinggi dalam mencapai tujuan organisasi. Dengan kata lain insentif kerja adalah uang disamping gaji yang dibayarkan oleh pimpinan organisasi menurut (Sahir &

Panjaitan, 2019) untuk mengevaluasi prestasi kerja dan kontribusi karyawan terhadap organisasi.

Insentif adalah daya tarik yang disengaja yang diberikan oleh perusahaan, manajer dan karyawan. Tujuannya untuk ikut serta dalam membangun, memelihara dan memperkuat ekspektasi manajer dan karyawan agar memiliki semangat yang lebih besar untuk berpartisipasi di perusahaan dalam hal meningkatkan efisiensi kerja.

Oleh karena itu, tindakan insentif akan mendorong peningkatan prestasi kerja karyawan yang sangat membantu pencapaian tujuan perusahaan, pada akhirnya tindakan insentif akan membawa manfaat bagi kedua belah pihak, yaitu meningkatkan prestasi kerja bagi karyawan berarti meningkatkan pendapatan.

#### **2.1.2.6. *Waikato Environment for Knowledge Analysis (WEKA)***

*Waikato Environment for Knowledge Analysis* atau WEKA adalah seperangkat alat pembelajaran mesin praktis. *Waikato Environment for Knowledge Analysis (WEKA)* dikembangkan oleh Universitas Waikato di Selandia Baru untuk penelitian, pendidikan, dan berbagai aplikasi (Pujiono, Amborowati, & Suyanto, 2013). WEKA adalah software open source, jadi siapapun bisa menambahkan software ini (Haryati, Sudarsono, & Suryana, 2015).

WEKA dapat digunakan untuk memecahkan masalah data mining di dunia nyata, khususnya klasifikasi dengan menggunakan metode pembelajaran mesin. WEKA mencakup pemrosesan awal data, klasifikasi, regresi, pengelompokan, aturan asosiasi, dan alat visualisasi. Pengujian dengan WEKA lebih fleksibel dan



efektif karena pohon keputusan yang terbentuk hampir sama dengan perhitungan manual (Azwanti, 2018).



**Gambar 2.5** Tampilan Utama WEKA

## 2.2. Penelitian Terdahulu

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

1. Dari penelitian “Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5” yang dilakukan oleh (Sri Lestari & Saefudin, 2015) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membangun sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5 dengan kriteria-kriteria yang telah ditentukan oleh perusahaan.

2. Dari penelitian “Komparasi Algoritma C4.5 Dan *Naive Bayes* Yang diKembangkan Menjadi *Web Intelligence* Pada Perhitungan Bonus Tahunan Karyawan di PT.ABC” yang dilakukan oleh (Ibnu Alfarobi & Taransa Agasya Tutupoly, 2019) didapatkan informasi bahwa bonus karyawan dari kedua algoritma *Decission Tree* C4.5 dan *Naive Bayes* yang dimana memiliki nilai *Area Under Curve (AUC)* sebesar 0.991 dan *Accuracy* atau *Confusion Matrix* sebesar 99%.
3. Dari penelitian “Penerapan Algoritma C4.5 Dalam Analisa Kelayakan Penerima Bonus Tahunan Pegawai (Studi Kasus : PT. Multi Pratama Nauli Medan)” yang dilakukan oleh (Syahputra, M.Kom, 2015) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membantu seorang pimpinan perusahaan dalam menganalisa pegawai yang layak untuk menerima bonus tahunan.
4. Dari penelitian “Penerapan Algoritma C4.5 Pada Analisis Penentuan Kelayakan Penerima Bonus Tahunan Pegawai Bank Sinarmas Multifinance” yang dilakukan oleh (Intan Komala Dewi Patwari & Baiq Andriskha Candra Permana, 2019) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti dapat membantu seorang pemimpin dalam mengambil suatu keputusan apakah seorang pegawai layak atau tidak untuk diberi bonus.
5. Dari penelitian “Prediksi Waktu Kuliah Dengan Algoritma Classifier” yang dilakukan oleh (Rika Harman & Evan Rosiska, 2020) didapatkan informasi bahwa dengan penggunaan jenis algoritma C4.5 terbukti

dapat membantu mahasiswa mencari akar permasalahan dan menentukan penyebab utama mahasiswa tidak dapat menyelesaikan masa studinya tepat waktu.

6. 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 penyelarasan data dengan kategori atau kriteria dalam melakukan *data mining* sehingga waktu yang dibutuhkan untuk mendapatkan keputusan akhir menjadi lebih singkat.
7. Dari penelitian “*Performance improvement of datamining ini WEKA through GPU acceleration*” yang dilakukan oleh (Augusto & Char, 2014) didapatkan informasi bahwa proses yang memakan waktu banyak dapat mengadaptasi terhadap *GPU*. *show that set of operations that are time consuming can easily adapted to GPUs*. Ketika memparalelkan metode penggandaan matriks, metode ini dapat beradaptasi untuk memanfaatkan *CPU* dan *GPU* sesuai dengan ukuran matriks, metode ini mempercepat proses hingga 49%.
8. Dari penelitian “Penerapan Algoritma C.45 Untuk Kelayakan Penerima Uang Dan Beras” yang dilakukan oleh (Harman, 2018) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 diharapkan dapat memperbaiki kebijakan-kebijakan yang selama ini dianggap tidak adil atau faktor kedekatan lebih berperan besar untuk penerima bantuan untuk uang dan beras kedepannya setelah penelitian ini selesai dilakukan.

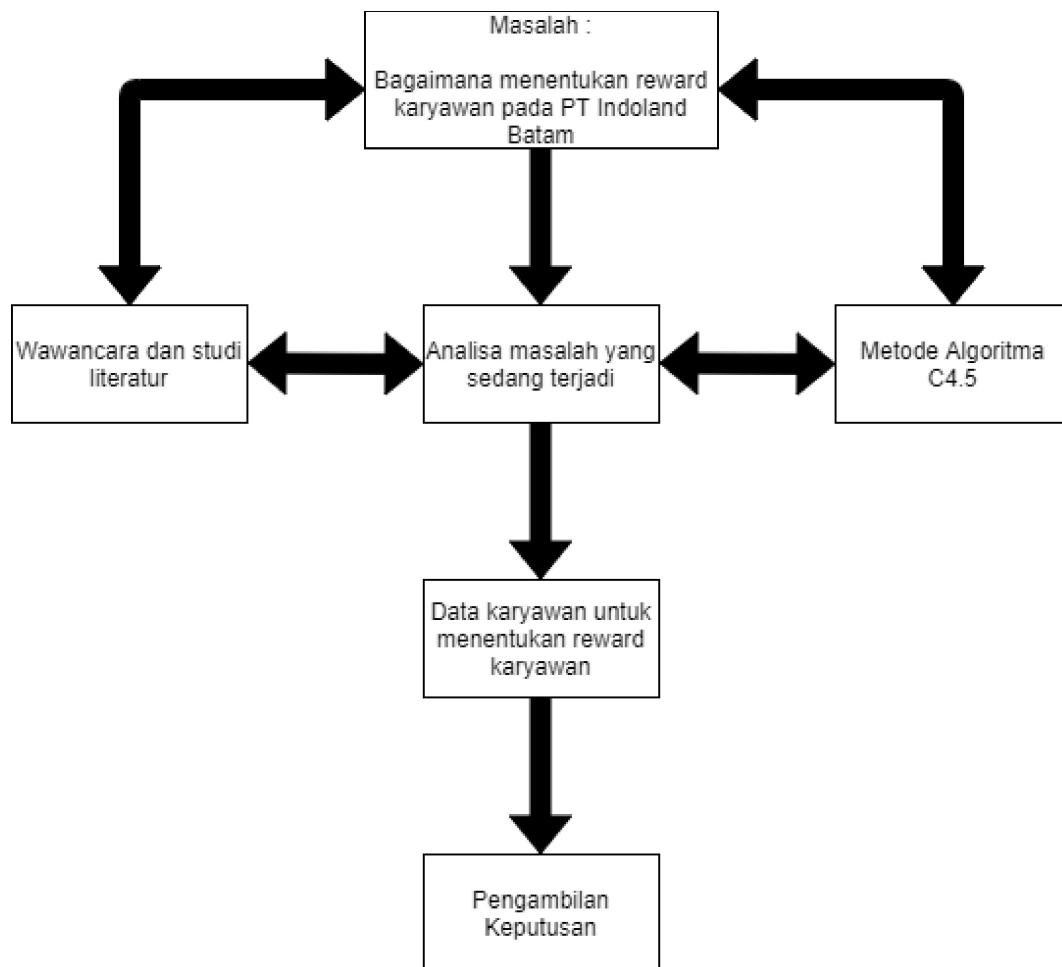
9. Dari penelitian “Analisis Dengan Metode Klasifikasi Menggunakan *Decision Tree* Untuk Prediksi Penentuan Resiko Kredit Pada Bank Bukopin Batam” yang dilakukan oleh (Intan Utma Sari, 2015) didapatkan informasi bahwa penggunaan jenis algoritma dari C4.5 diharapkan dapat membantu pihak bank agar lebih cepat dan mudah dalam memprediksi data dan menarik suatu kesimpulan dari data yang ada. Salah satu cara memprediksi data tersebut dengan menggunakan *software Dtree* dan *Decision Tree*. *Dtree* dapat menghasilkan pohon keputusan, salah satu nya yaitu hasil keputusan resiko dari jumlah kredit pemilikan rumah berdasarkan jumlah nasabah sehingga dan dengan teknik *Decision Tree* diharapkan dapat membantu pihak bank agar lebih cepat dan mudah dalam memprediksi data dan menarik suatu kesimpulan dari data yang ada.
10. Penelitian yang berjudul “implementation of decision trees algorithm C4.5” yang dilakukan oleh (Chauhan & Chauhan, 2013) menunjukkan bahwa algoritma C4.5 menunjukkan hasil yang lebih akurat dan pengimplementasian pohon keputusan dengan metode algoritma C4.5 sangat mudah dan cepat. Algoritma ini lebih cocok untuk data sets yang mengandung tidak terlalu banyak data.
11. Penelitian yang berjudul “Performance evaluation of the GIS-based datamining techniques of best-first decision tree, random forest, and naïve bayes tree for landslide susceptibility modeling” yang dilakukan

oleh (Chen, Zhang, Li, & Shahabi, 2018) menghasilkan bahwa random forest model merupakan yang paling akurat prediksinya.

12. Dari penelitian "Reservoir Inflow Forecasting Using ID3 and C4.5 Decision Tree Model" yang dilakukan oleh (Charoenporn, 2017) didapatkan informasi<sup>29</sup> 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.

### **2.3. Kerangka Pemikiran**

Berikut adalah kerangka pemikiran dalam penelitian ini, yang antara lain dapat digambarkan sebagai berikut:



**Gambar 2.6** Kerangka Berpikir

**Sumber:** (Penulis, 2020)