

## **BAB 1**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Toko mobil Serasi Motor adalah salah satu Toko yang sudah memiliki tiga cabang yang bergerak dalam penjualan alat-alat mobil. Toko mobil ini selalu memproduksi stok barang setiap harinya dengan contoh barang-barang *spare part*. Stok barang merupakan salah satu faktor penting yang harus diperhatikan oleh setiap Toko. Jumlah barang masuk dan barang keluar perlu diperhatikan, tujuannya untuk menjaga stok barang yang tersedia di gudang menjadi stabil. Stok barang stabil yang dimaksud dalam gudang adalah barang tidak kelebihan stok maupun barang tidak kehabisan stok. Karena permasalahan yang dihadapi pada penjualan barang adalah setiap barang memiliki minat beli berbeda-beda. Barang dengan minat beli yang tinggi akan cepat habis dibandingkan dengan barang yang memiliki minat beli rendah.

Persediaan barang di gudang harus dijaga setiap harinya, karena jumlah penjualan yang memiliki minat beli tinggi mengakibatkan stok barang yang tersedia tidak stabil dan dapat berdampak langsung kepada konsumen. Kepuasan konsumen pada toko dapat berkurang karena tidak tersedianya barang yang hendak dibeli. Keterersediaan barang yang tidak dikelola dengan baik juga berdampak pada toko, misalkan barang habis pada saat permintaan konsumen tinggi maka yang akan terjadi adalah permintaan barang harus diundur ataupun dibatalkan.

Permasalahan lainnya adalah penimbunan stok barang yang berlebihan. Kesalahan prediksi penjualan menjadi salah satu alasan yang membuat toko memesan stok barang dalam jumlah yang besar yang akhirnya tidak habis terjual sehingga stok barang tersebut membengkak di gudang. Penimbunan ini menyebabkan toko merugi dikarenakan dana yang harus keluar untuk proses penyimpanan barang. Selain itu juga terdapat istilah *backorder*, dimana barang yang sudah dipesan konsumen belum tersedia, sehingga konsumen harus menunggu sampai barang tersebut tersedia.

Permasalahan-permasalahan yang terjadi diatas disebabkan karena kesulitan menentukan stok minimum tiap hari yang harus dipenuhi berdasarkan minat konsumen. Penentuan stok barang dapat ditentukan dari berbagai aspek. Pada penelitian ini data yang akan digunakan untuk menentukan waktu pembelian stok barang adalah data *backorder*. Dari data tersebut akan dilihat faktor yang mempengaruhi *backorder* sehingga toko dapat memperkirakan stok barang sebelum dipesan oleh konsumen.

Salah satu metode yang bisa digunakan untuk penentuan stok barang adalah *Algoritma C4.5*. merupakan *Algoritma* klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari *Algoritma* yang lainnya. Kelebihan *Algoritma C4.5* dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik. *Algoritma C4.5* digunakan untuk melakukan klasifikasi dan segmentasi yang banyak dimanfaatkan untuk melakukan prediksi terhadap suatu

kasus dengan proses menentukan model atau fungsi untuk menggambarkan *class* atau konsep dari suatu data. Klasifikasi (Han, 2006) adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui [1].

Pada penelitian terdahulu, *Algoritma C4.5* dapat membantu dalam menentukan mahasiswa yang diterima dan ditolak sebagai penerima beasiswa dari STT Harapan Medan. *Algoritma C4.5* memiliki kelebihan utama yaitu dapat menghasilkan model berupa *tree* atau aturan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, dapat menangani atribut bertipe diskrit dan numerik. Pada *Algoritma C4.5*, model yang dihasilkan oleh proses “belajar” dari data pelatihan berupa sebuah *decision tree*. *Decision tree* ini kemudian dapat dimanfaatkan untuk memprediksi kelas dari kasus yang baru (Rismayanti.2016) [2].

Pada penelitian selanjutnya *Algoritma C4.5* digunakan untuk menentukan kelayakan penerima bonus tahunan pegawai pada PT. Multi Pratama Nauli Medan. Salah satunya dalam meningkatkan kualitas dan loyalitas pegawai pada PT. Multi Pratama Nauli Medan, maka setiap tahunnya pimpinan perusahaan memberikan bonus pada pegawai yang berprestasi dengan kriteria-kriteria yang telah ditetapkan. Karena banyaknya kriteria dan data pegawai yang dianalisa, maka seorang pimpinan perusahaan harus bekerja keras dalam menganalisa pegawai mana yang layak dalam menerima bonus tahunan ini. Maka diperlukan pengolahan data lebih lanjut untuk menemukan informasi atau pengetahuan baru

yang berguna sebagai pembantu dalam mengambil keputusan. Oleh karena itu, pembentukan suatu pola keputusan sangat diperlukan karena dapat menghindari permasalahan tersebut. Penerapan *Algoritma C4.5* merupakan salah satu solusi pemecahan kasus yang sering digunakan dalam pemecahan masalah pada teknik klasifikasi yang memiliki karakteristik yaitu dengan proses penentuan nilai *Entropy* dan nilai *Gain* dari kemungkinan setiap kriteria yang menjadi acuan keputusan yang dilanjutkan dengan proses perangkaian dari hasil keputusan. Keluaran dari *Algoritma C4.5* yaitu berupa sebuah pohon keputusan (*decision tree*). Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk mengubah data menjadi pohon keputusan yang akan menghasilkan aturan-aturan keputusan (Guntur. 2014) [3].

Dari uraian di atas, penulis mengangkat judul penelitian “Penerapan *Algoritma C4.5* Untuk Penentuan Stok Barang Alat-alat Mobil.”

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dikemukakan, yang menjadi pokok permasalahan adalah bagaimana membangun aplikasi penentuan stok barang alat-alat mobil menggunakan metode *Data Mining Algoritma C4.5*?

## **1.3. Tujuan Penelitian**

Adapun tujuan dari penelitian ini adalah untuk membuat aplikasi penentuan stok barang alat-alat mobil dalam membantu kinerja dalam pengecekan stok barang menggunakan metode *Algoritma C4.5*.

#### **1.4. Batasan Masalah**

Dalam penelitian ini, peneliti memberikan beberapa batasan masalah yaitu:

1. Metode yang digunakan dalam penentuan stok barang alat-alat mobil yaitu metode *Algoritma C4.5*.
2. Aplikasi penentuan stok barang alat-alat mobil meliputi: pemasukan, pengeluaran, jumlah stok barang, serta penjualan yang laris.
3. Aplikasi penentuan stok barang alat-alat mobil dikembangkan dengan menggunakan bahasa pemrograman *PHP* dan *MySQL*.

#### **1.5. Manfaat Penelitian**

Manfaat dari penelitian ini adalah aplikasi dapat membantu dalam kinerja dan penentuan stok barang di gudang apakah barang tersebut perlu di tambah stoknya atau tidak.

#### **1.6. Metodologi Penelitian**

Tahapan-tahapan yang akan dilakukan pada pelaksanaan penelitian adalah sebagai berikut:

##### **1. Studi Literatur**

Pada tahapan ini peneliti mengumpulkan bahan referensi berkaitan dengan *Data Mining Algoritma C4.5* dalam penentuan stok barang, dan memahami bahasa pemrograman *Php* dan *MySQL* dari berbagai jurnal, skripsi, buku, artikel, dan berbagai sumber referensi lainnya.

##### **2. Analisa Masalah**

Pada tahapan ini dilakukan analisis untuk setiap informasi yang telah diperoleh dari analisa lapangan secara langsung maupun tidak

langsung serta melakukan wawancara secara langsung dengan pihak-pihak terkait studi kasus agar mendapatkan pemahaman akan masalah yang dihadapi dan metode yang akan digunakan untuk menyelesaikan permasalahan.

### 3. Perancangan Sistem

Pada tahapan ini dilakukan perancangan sistem sesuai dengan hasil dari tahap sebelumnya.

### 4. Implementasi

Pada tahapan ini hasil dari analisis dan perancangan sistem akan di implementasikan ke dalam kode program.

### 5. Pengujian

Pada tahapan ini dilakukan pengujian terhadap aplikasi penerapan stok barang dengan metode *Algoritma C4.5* dalam membantu kinerja dan pengecekan stok barang dapat memberikan hasil yang baik.

### 6. Dokumentasi dan Penyusunan Laporan

Pada tahap terakhir membuat dokumentasi dan menyusun laporan hasil dari analisi dan implementasi dari penelitian tersebut.

## **1.7. Sistematika Penulisan**

Sistematika penulisan dari skripsi ini terdiri dari lima bagian utama sebagai berikut :

## **BAB 1 PENDAHULUAN**

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

## **BAB 2 LANDASAN TEORI**

Bab ini berisi teori-teori yang digunakan pada penelitian ini. Teori-teori yang berhubungan dengan aplikasi *Data Mining Algoritma C4.5* dalam penentuan stok barang.

## **BAB 3 METODE PENELITIAN**

Bab ini berisi tahapan-tahapan dalam pengumpulan data, perancangan sistem, perumusan masalah dan analisa.

## **BAB 4 ANALISA DAN PERANCANGAN**

Bab ini berisi analisa dan perancangan aplikasi penerapan metode *Algoritma C4.5*. untuk membangun aplikasi penentuan stok barang alat-alat mobil.

## **BAB 5 IMPLEMENTASI DAN PENGUJIAN**

Bab ini berisi implementasi dari analisa dan perancangan dan pengujian pada aplikasi yang berhasil dibangun.

## **BAB 6 PENUTUP**

Bab ini berisi rangkuman dari hasil penelitian yang telah dilakukan dan saran-saran untuk pengembangan aplikasi atau penelitian selanjutnya.

## **BAB 2**

### **LANDASAN TEORI**

#### **2.1 Serasi Motor**

Serasi Motor merupakan sebuah unit dagang yang bergerak dibidang penjualan barang alat-alat mobil. Dan letaknya berada di desa Talikumain. Dimana alat-alat mobil ini setiap harinya harus memenuhi kebutuhan konsumen sehingga penjualan merupakan hal yang utama dalam perusahaan/toko tersebut. Dalam menangani penjualan alat-alat mobil tersebut, Serasi Motor masih menggunakan cara yang konvensional/manual dengan cara melakukan pencatatan setiap transaksi dengan menggunakan buku besar, dan Toko ini sudah memiliki tiga cabang yang bergerak dalam penjualan alat-alat mobil.

#### **2.2 Persediaan / Stok barang**

Pengertian persediaan, adalah suatu aktiva yang meliputi barang-barang milik perusahaan dengan maksud untuk dijual dalam suatu periode usaha yang normal, atau persediaan barang-barang yang masih dalam pengerjaan/proses produksi, ataupun persediaan bahan baku yang menunggu penggunaannya dalam suatu proses produksi.

#### **2.3 Data Mining**

*Data Mining* adalah gabungan sejumlah disiplin ilmu komputer yang mendefinisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence, machine learning, statistics, dan database systems* [4].

*Data Mining* adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [5].

*Data Mining* atau penambahan kata adalah teknik yang relatif cepat dan mudah untuk menemukan pola atau relasi antar data, secara otomatis. Dengan menggabungkan empat disiplin ilmu komputer seperti definisi diatas, pengetahuan bisa ditemukan dalam lima proses berurutan : seleksi, pemrosesan, transformasi, *Data Mining*, dan interpretasi/evaluasi.

*Data Mining* adalah metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode *Data Mining* ini dapat digunakan untuk mengambil keputusan dimasa depan. *Data Mining* ini juga dikenal dengan istilah *pattern recognition*. *Data Mining* merupakan metode pengolahan data berskala besar oleh karena itu *Data Mining* ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian *Data Mining* membahas metode-metode seperti *clustering*, *klasifikasi*, *regresi*, *seleksi variable* dan *market basket analisis*.

Dari definisi-definisi diatas, maka dapat disimpulkan bahwa secara umum *Data Mining* adalah suatu kegiatan analisa data untuk mencari suatu pola tertentu, dengan jumlah data yang besar dan bertujuan untuk menghasilkan informasi yang dapat digunakan dan dikembangkan lebih lanjut.

### 2.1.1 Tahap-tahap *Data Mining*

*Data Mining* merupakan inti dari proses *Knowledge Discovery Database* (KDD), meliputi dugaan *Algoritma* yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui. KDD merupakan penyelesaian masalah dengan menganalisa data yang ada pada *database* dengan data tersimpan secara elektronik dan pencariannya dilakukan otomatis seperti pada komputer.

1. *Data Selection* Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *Data Mining*, disimpan suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/Cleaning* sebelum proses *Data Mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang *inkonsisten*, memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *Data Mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode,

atau *Algoritma* dalam *Data Mining* sangat bervariasi pemilihan metode atau *Algoritma* yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation/ Evaluation* Pola informasi yang dihasilkan dari proses *Data Mining*, perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada.

Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna.

#### 2.1.2 Teknik-teknik *Data Mining*

Dengan definisi *Data Mining* yang luas, ada banyak jenis teknik analisa yang dapat digolongkan dalam *Data Mining*. Karena keterbatasan tempat, disini penulis akan memberikan sedikit gambaran tentang tiga teknik *Data Mining* yang paling populer.

##### a. *Association Rule Mining*

*Association rules* (aturan asosiasi) atau *affinity analysis* (analisis afinitas) berkenaan dengan studi tentang “apa bersama apa”. Sebagai contoh dapat berupa berupa studi transaksi di supermarket, misalnya seseorang yang membeli susu bayi juga membeli sabun mandi. Pada kasus ini berarti susu bayi bersama dengan sabun mandi. Karena awalnya berasal dari studi tentang *database* transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama produk apa, maka aturan asosiasi juga sering di namakan *market basket analysis*.

“*if-then*” atau “jika-maka”. Aturan ini dihitung dari data yang sifatnya *probabilistic*.

Analisis asosiasi dikenal juga sebagai salah satu metode *Data Mining* yang menjadi dasar dari berbagai metode *Data Mining* lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola *frekuensi* tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan *Algoritma* yang efisien. Penting tidaknya suatu aturan *assosiatif* dapat diketahui dengan dua *parameter, support* (nilai penunjang) yaitu prosentase kombinasi item tersebut dalam *database* dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antara item dalam aturan *assosiatif*. *Analisis asosiasi* didefinisikan suatu proses untuk menemukan semua aturan *assosiatif* yang memenuhi syarat minimum untuk *support (minimum support)* dan syarat minimum untuk *confidence*.

#### b. *Classification*

Dalam klasifikasi, terdapat *target variable* kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

#### C. *Clustering*

*Clustering* termasuk metode yang sudah cukup dikenal dan banyak dipakai dalam *Data Mining*. Sampai sekarang para ilmuwan dalam bidang *Data Mining* masih melakukan berbagai usaha untuk melakukan perbaikan model *clustering* karena metode yang dikembangkan sekarang masih bersifat *heuristic*. Usaha-usaha untuk menghitung jumlah *cluster* yang optimal dan pengklasteran yang paling baik masih terus dilakukan.

Dengan demikian menggunakan metode yang sekarang, tidak bisa menjamin hasil pengklasteran sudah merupakan hasil yang optimal. Namun, hasil yang dicapai biasanya sudah cukup bagus dari segi praktis. Tujuan utama dari metode *clustering* adalah pengelompokan sejumlah data/objek ke dalam *cluster* (grup) sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin. Dalam *clustering* metode ini berusaha untuk menempatkan obyek yang mirip (jaraknya dekat) dalam satu *cluster* dan membuat jarak antar *cluster* sejauh mungkin. Ini berarti obyek dalam satu *cluster* sangat mirip satu sama lain dan berbeda dengan obyek dalam *cluster-cluster* yang lain. Dalam metode ini tidak diketahui sebelumnya berapa jumlah *cluster* dan bagaimana pengelompokannya.

Berikut ini ada 9 (Sembilan) *Algoritma* penggalian data yang paling populer berdasarkan konferensi ICDM '06:

1. *C4.5*
2. *k-Means*
3. *SVM 24*
4. *Apriori*
5. *EM*
6. *PageRank*
7. *AdaBoost*
8. *KNN*
9. *Naive Bayes*

## 2.4 *Algoritma C4.5*

*Algoritma C4.5* merupakan salah satu *Algoritma* yang digunakan untuk membentuk pohon keputusan. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu [6].

*Algoritma C4.5* merupakan *Algoritma* klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari *Algoritma* yang lainnya. Kelebihan *Algoritma C4.5* dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik. Dalam mengkonstruksi pohon, *Algoritma C4.5* membaca seluruh sampel *data training* dari *storage* dan memuatnya ke memori. *Algoritma* Pohon Keputusan *C4.5* atau *Classification version 4.5* adalah pengembangan dari *Algoritma ID3*. Oleh karena pengembangan tersebut, *Algoritma C4.5* mempunyai prinsip dasar kerja yang sama dengan *Algoritma ID3*. Secara umum, proses *Algoritma C4.5* untuk membangun pohon keputusan adalah sebagai berikut.

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang

4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama secara khusus, *Algoritma C4.5 Decision Tree* menggunakan kriteria *split* yang telah dimodifikasi yang dinamakan *Gain Ratio* dalam proses pemilihan *split* atribut. *Split* atribut merupakan proses utama dalam pembentukan pohon keputusan (*Decision Tree*) di *C4.5*.

Tahapan dari *Algoritma C4.5* adalah sebagai berikut.

- a. Menghitung nilai *Entropy*,
- b. Menghitung nilai *Gain* untuk masing-masing atribut,
- c. Atribut yang memiliki *Gain* tertinggi dipilih menjadi akar (*root*) dan atribut yang memiliki nilai *Gain* lebih rendah dari akar (*root*) dipilih menjadi cabang (*branches*),
- d. Menghitung lagi nilai *Gain* tiap-tiap atribut dengan tidak mengikutsertakan atribut yang terpilih menjadi akar (*root*) ditahap sebelumnya,
- e. Atribut yang memiliki *Gain* tertinggi dipilih menjadi cabang (*branches*),
- f. Mengulangi langkah ke-4 dan ke-5 sampai dengan dihasilkan nilai *Gain* = 0 untuk semua atribut yang tersisa.

Dalam memilih sebuah atribut menjadi akar, dilakukan perhitungan nilai dari atribut yang ada. Nilai *Gain* yang paling tinggi dijadikan *root* pada pohon keputusan. Untuk menghitung nilai *Gain* di gunakan rumus:

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

Keterangan:

S : Himpunan kasus,

A : Atribut,

n : Jumlah partisi dalam atribut,

|Si| : Jumlah kasus pada partisi ke-i,

|S| : Jumlah kasus

Sehingga akan diperoleh nilai *Gain* dari atribut yang paling tertinggi. *Gain* adalah salah satu *attribute selection measure* yang digunakan untuk memilih *test attribute* tiap *node* pada *tree*. Atribut dengan *information Gain* tertinggi dipilih sebagai *test attribute* dari suatu *node*. Sedangkan untuk perhitungan nilai *Entropy* dapat dilakukan dengan rumus:

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

Keterangan:

S : Himpunan kasus,

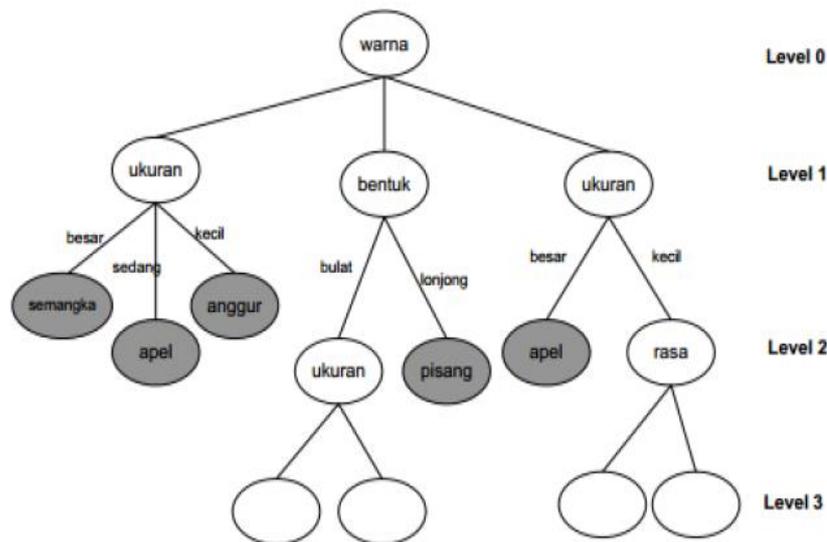
n : Jumlah partisi dalam atribut,

pi: Proporsi dari Si terhadap S

## 2.5 Pohon keputusan (*decision tree*)

Dalam *decision tree* tidak menggunakan vector jarak untuk mengklasifikasikan obyek. sering kali data observasi mempunyai atribut-atribut yang bernilai nominal. Seperti yang diilustrasikan pada gambar 2.1 misalkan obyeknya adalah sekumpulan buah-buahan yang bisa di bedakan berdasarkan atribut bentuk, warna, ukuran dan rasa. Bentuk, warna, ukuran dan rasa adalah besaran nominal, yaitu bersifat kategoris dan tiap nilai tidak bisa di jumlahkan atau dikurangkan. Dalam atribut warna ada beberapa nilai yang mungkin yaitu hijau, kuning, merah. Dalam atribut ukuran ada nilai besar, sedang dan kecil.

Dengan nilai-nilai atribut ini, kemudian dibuat *decision tree* untuk menentukan suatu obyek termasuk jenis buah apa jika nilai tiap-tiap atribut di berikan [7].



**Gambar 2.1** *decision tree*

Ada beberapa macam *Algoritma decision tree* diantaranya *CART* dan *C4.5*. Beberapa isu utama dalam *decision tree* yang menjadi perhatian yaitu seberapa detail dalam mengembangkan *decision tree*, bagaimana mengatasi atribut yang bernilai *continues*, memilih ukuran yang cocok untuk penentuan atribut, menangani data *training* yang mempunyai data yang atributnya tidak mempunyai nilai, memperbaiki efisiensi perhitungan [7].

*Decision tree* sesuai digunakan untuk kasus-kasus yang keluarannya bernilai diskrit. Walaupun banyak variasi model *decision tree* dengan tingkat kemampuan dan syarat yang berbeda, pada umumnya beberapa ciri yang cocok untuk diterapkannya *decision tree* adalah sebagai berikut :

1. Data dinyatakan dengan pasangan atribut dan nilainya
2. Label/keluaran data biasanya bernilai diskrit

3. Data mempunyai *missing value* (nilai dari suatu atribut tidak diketahui)

Dengan cara ini akan mudah mengelompokkan obyek ke dalam beberapa kelompok. Untuk membuat *decision tree* perlu memperhatikan hal-hal berikut ini :

1. Atribut mana yang akan dipilih untuk pemisahan obyek
2. Urutan atribut mana yang akan dipilih terlebih dahulu
3. Struktur *tree*
4. Kriteria pemberhentian
5. *Pruning*

*Decision tree* adalah struktur *flowchart* yang mempunyai *tree* (pohon), dimana setiap simpul internal menandakan suatu tes atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun merepresentasikan kelas atau distribusi kelas.

## **2.6 Website**

Aplikasi *web* adalah suatu aplikasi yang diserahkan melalui *web* yang mengombinasikan karakteristik dari hipermedia *web* dan aplikasi perangkat lunak *web* [8].

## **2.7 Hypertext Proprecessor**

Hypertext Proprecessor adalah akronim dari Hypertext Preprocessor, yaitu suatu bahasa pemrograman berbasis kode-kode (*script*) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke *web browser* menjadi kode HTML [9].

## 2.8 MySQL

*MySQL* adalah sebuah basis data yang mengandung satu atau jumlah tabel. Tabel terdiri atas sejumlah baris dan setiap baris mengandung satu atau sejumlah tabel. Tabel terdiri atas sejumlah baris dan setiap baris mengandung satu atau sejumlah tabel". *MySQL* adalah *database server open source* yang cukup populer keberadaannya. Dengan berbagai keunggulan yang dimiliki, membuat *software database* ini banyak di gunakan oleh praktisi untuk membangun suatu *project*. Adanya fasilitas API (*Application Programming Interface*) yang dimiliki oleh *MySQL*, memungkinkan bermacam-macam aplikasi komputer yang di tulis dengan berbagai bahasa pemrograman dapat mengakses basis data *MySQL*. Tipe data *MySQL*, Tipe data *MySQL* adalah data yang terdapat dalam sebuah *table* berupa *field-field* yang berisi nilai dari data tersebut. Nilai data dalam *field* memiliki tipe sendiri-sendiri [10].

## 2.9 Data Flow Diagram

*Data Flow Diagram* adalah suatu model logika data atau proses yang dibuat untuk menggambarkan dari mana asal data atau kemana tujuan data yang keluar dari sistem, dimana data disimpan, proses apa yang menghasilkan data tersebut dan interaksi antara data yang disimpan dan proses yang dikenakan pada data tersebut [11].

## 2.10 Entity Relationship Diagram

*Entity Relationship Diagram (Diagram E-R)* adalah yang digunakan untuk menggambarkan model *Entity Relationship* yang berisi komponen-komponen. himpunan entitas dan himpunan relasi yang masing-masing dilengkapi dengan

atribut-atribut yang mempersentasikan seluruh fakta dari dunia nyata yang kita tinjau. LRS (*Logical Record Structure*) merupakan representasi dari struktur *record-record* pada tabel-tabel yang terbentuk dari hasil antar himpunan entitas [12].

### 2.11 Penulisan Terkait

No	Penulis / Tahun	Judul	Metode	Hasil
1	( G L Pritalia, 2018)	Penerapan <i>Algoritma C4.5</i> untuk Penentuan Ketersediaan Barang <i>E-commerce</i>	<i>Data Mining</i> , <i>Algoritma C4.5</i>	<i>Data Mining</i> dengan <i>Algoritma C4.5</i> dapat diimplementasikan untuk penentuan ketersediaan barang <i>e-commerce</i> dengan dua kategori yaitu barang tersebut dalam kondisi tidak atau belum tersedia untuk pemesanan barang untuk pelanggan/ <i>backorder</i> dan barang dalam kondisi tersedia untuk pemesanan barang untuk pelanggan.
2	(Pandu Pratama Putra, Andi Supriadi Chan 2018)	Pengembangan aplikasi Prediksi Stock Motor Menggunakan <i>Algoritma C4.5</i> Sebagai Bagian dari Sistem Pengambilan Keputusan	<i>Data Mining</i> , <i>Algoritma C4.5</i>	Berdasarkan proses perhitungan yang telah dilakukan melalui beberapa langkah pada proses penerapan <i>Data Mining</i> menggunakan <i>Algoritma C4.5</i> maka diperoleh pohon keputusan tersebut

## BAB 3

### METODE PENELITIAN

Penelitian ini dilakukan dengan melaksanakan tahapan demi tahapan yang berhubungan. Tahapan-tahapan tersebut dijabarkan dalam metode penelitian. Metode penelitian diuraikan kedalam bentuk skema yang jelas, teratur, dan sistematis. Berikut tahapan-tahapan penelitian dilihat pada diagram alir 3.1 di bawah:



**Gambar 3.1 Tahapan Metodologi Penelitian**

Dalam metodologi penelitian dijabarkan tahapan-tahapan yang dilakukan dalam penelitian. Metodologi penelitian terdiri dari beberapa tahapan yang terkait secara sistematis. Tahapan ini diperlukan untuk memudahkan dalam melakukan penelitian. Tahapan yang dilakukan dalam penelitian adalah sebagai berikut :

### **3.1 Pengumpulan Data**

Pada tahap ini di lakukan pengumpulan data yang berhubungan dengan penelitian dan pembuatan sistem, yaitu dengan :

#### **1. Wawancara (*Interview*)**

Melakukan wawancara secara langsung kepada pihak yang bertanggung jawab atas Toko penjualan barang atau Toko Serasi Motor.

#### **2. Studi Kepustakaan**

Studi kepustakaan dilakukan dengan cara mempelajari buku-buku, jurnal-jurnal dan artikel-artikel di internet yang berhubungan dengan permasalahan yang dibahas.

### **3.2 Perumusan Masalah**

Berdasarkan hasil dari tahapan yang dijelaskan sebelumnya yang telah dijabarkan, maka tahapan selanjutnya adalah tahapan perumusan masalah. Pada tahapan perumusan masalah akan dirumuskan masalah yang dianggap sebagai penelitian dalam tugas akhir ini. Permasalahan-permasalahan yang menjadi rumusan masalah dalam penelitian ini didapatkan dari penelitian lapangan secara langsung serta melakukan wawancara terhadap objek terkait dengan penelitian secara langsung. Solusi yang didapatkan pada tahapan perumusan masalah ini

yang akan menjadi judul penelitian tugas akhir ini yaitu dengan judul “ Penerapan *Algoritma C4.5* untuk penentuan stok barang alat-alat mobil”.

### **3.3 Analisa Sistem**

Setelah menganalisa sistem lama, maka tahapan dapat dilanjutkan dengan menganalisa sistem yang baru. Analisa dalam pembuatan sistem ini menggunakan metode *Algoritma C4.5* serta menganalisa kebutuhan sistem. Data-data yang dibutuhkan untuk memulai pembuatan sistem ini dimasukkan ke dalam analisa data sistem untuk menentukan berapa jumlah stok barang.

### **3.4 Perancangan Sistem**

Setelah melakukan analisa, maka kemudian dilanjutkan dengan perancangan sistem berdasarkan analisa permasalahan yang telah dilakukan sebelumnya.

#### **3.4.1 Perancangan Basis Data**

Setelah menganalisa sistem yang akan dibuat, maka tahap selanjutnya adalah analisa dan perancangan basis data yang dilakukan untuk melengkapi komponen sistem.

#### **3.4.2 Perancangan Struktur Menu**

Rancangan struktur menu diperlukan untuk memberikan gambaran terhadap menu-menu atau fitur pada sistem yang akan dibangun.

#### **3.4.3 Perancangan Antar Muka (*Interface*)**

Untuk mempermudah komunikasi antara sistem dengan pengguna, maka perlu dirancang antar muka (*interface*). Dalam perancangan *interface* hal

terpenting yang ditekankan adalah bagaimana menciptakan tampilan yang baik dan mudah dimengerti oleh pengguna.

### **3.5 Implementasi dan Pengujian**

#### **3.5.1 Implementasi Sistem**

Implementasi sistem merupakan suatu konversi dari desain sistem yang telah dirancang kedalam sebuah program komputer dengan menggunakan bahasa pemrograman *PHP* berbasis *website* dengan *database MySQL*. Adapun fungsi-fungsi perancangan aplikasi *Algoritma C4.5* ini adalah *Input* data stok barang, penyimpanan data stok barang, pengubahan jumlah data, penghapusan data, pengolahan data, pembuatan laporan yang dibutuhkan dan batasan wewenang atau otorisasi yang jelas kepada pemakai program aplikasi.

#### **3.5.2 Pengujian Sistem**

Pengujian sistem dilakukan dengan cara menggunakan *Blackbox* dan *User Acceptance Test*. Pada *Blackbox* pengujian program aplikasi *Data Mining Algoritma C4.5* ini berfokus pada perangkat lunak untuk mendapatkan serangkaian kondisi *input* yang seluruhnya menggunakan persyaratan fungsional dalam suatu program. Pengujian ini di uji cobakan kepada *user*, dan diberikan angket yang didalamnya berisi pertanyaan seputar tugas akhir ini.

### **3.6 Kesimpulan dan Saran**

Tahapan terakhir adalah menarik kesimpulan dari hasil penelitian yang didapatkan dalam penerapan metode *Algoritma C4.5* dalam penentuan stok barang. Pada tahapan ini juga berisikan saran penulis bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.