

**IMPLEMENTASI METODE *K-NEAREST NEIGHBOR*
DALAM PERAMALAN PENJUALAN
MOBIL BEKAS DI
KOTA BATAM**

SKRIPSI



Oleh :

M.Firmansyah Rafsyanzani s

140210322

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2021**

**IMPLEMENTASI METODE K-NEAREST NEIGHBOR
DALAM PERAMALAN PENJUALAN
MOBIL BEKAS DI
KOTA BATAM**

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana**



Oleh :

M.Firmansyah Rafsyanzani s

140210322

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2021**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Muhammad Firmansyah Rafsyanzani S

NPM : 140210322

Fakultas : Teknik dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

***IMPLEMENTASI METODE K-NEAREST NEIGHBOR DALAM PERAMALAN
PENJUALAN MOBIL BEKAS DI KOTA BATAM***

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 21 Januari 2021


METERAI
NPM
140210322
6000
ERATY PERKUMPAHAN
Muhammad Firmansyah Rafsyanzani S

**IMPLEMENTASI METODE K-NEAREST NEIGHBOR DALAM
PERAMALAN PENJUALAN MOBIL BEKAS
DI KOTA BATAM**

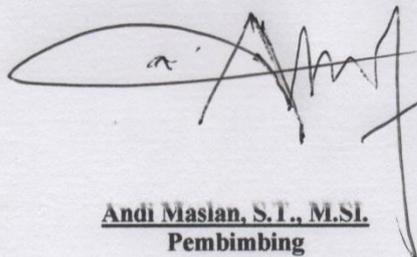
SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana**

**Oleh
Muhammad Firmansyah Rafsyanzani S
140210322**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, 27 Januari 2021



**Andi Maslan, S.T., M.Si.
Pembimbing**



Universitas Putera Batam

ABSTRAK

Perkembangan teknologi informasi mulai memasuki bagian terpenting dalam suatu bidang pekerjaan, informasi yang didapat dalam suatu pekerjaan menyangkut hal umum dan khusus. Komputer merupakan salah satu hasil dari perkembangan teknologi canggih yang dapat digunakan untuk membantu segala jenis pekerjaan termasuk transaksi pada pekerjaan bisnis dengan kemudahan yang ada didalam *Software* komputer seorang user dapat melakukan operasional transaksi dengan cepat. Penggunaan komputer juga dibutuhkan oleh seorang marketing. Penjualan mobil *PT. Auto Grace* adalah perusahaan yang bergerak dibidang jual-beli mobil bekas, berkedudukan di wilayah kota Batam. Penjualan mobil pada perusahaan *PT. Auto Grace* saat ini masih belum stabil dan mengakibatkan banyaknya mobil tidak terjual dan tidak terawat karena kurangnya minat konsumen untuk membeli. Agar perusahaan tidak mengalami masalah penjualan, maka pimpinan perusahaan harus mampu mencari solusi dengan cara mampu melakukan prediksi. Sehingga penelitian ini bertujuan untuk melakukan prediksi penjualan mobil menggunakan algoritma K-NN. Dataset yang digunakan untuk memprediksi di ambil dari data penjualan dari tahun 2017 sampai dengan 2020. Setelah dilakukan analisis menggunakan algoritma K-NN, dengan bantuan *Software Rapidminer*, maka diperoleh tingkat akurasi dalam memprediksi penjualan sebesar 63.89% dengan tingkat presisi sebesar 100%. Sehingga di ambil kesimpulan bahwa penjualan mobil pada tahun 2021 akan mengalami penurunan, karena dipengaruhi oleh mewabahnya virus covid-19 yang mengakibatkan terjadinya resesi ekonomi.

Kata Kunci : *Data Mining*, Penjualan Mobil, *K-Nearest Neighbour*, *Rapidminer Software*

ABSTRACT

The development of information technology begins to enter the most important part in a field of work, information obtained in a job is general and specific. Computers are one of the results of technological developments that can be used to help all types of work, including transactions in business work, with those in the computer Software, a user can perform transaction operations quickly. The use of computers is also needed by a marketing. Car sales of PT. Auto Grace is a company engaged in the sale and purchase of used cars, an area domiciled in the city of Batam. Car sales at PT. Auto Grace is currently still unstable and many cars are not sold and are not maintained due to the lack of consumer interest in buying. So that the company does not experience sales problems, the company leadership must be able to find solutions in a way that is able to predict. So this study aims to predict car sales using the K-NN algorithm. The data set used to predict is taken from sales data from 2017 to 2020. After analysis using the K-NN algorithm, with the help of Rapidminer Software, the accuracy rate in predicting sales is 63.89% with a precision level of 100%. . The question of take that sale car at year 2021 will experience a decline, due to the outbreak of the Covid-19 virus which has resulted in an economic recession.

Keyword: *Data Mining, Car sale, K-Nearest Neighbour, Rapidminer Software*

KATA PENGANTAR

Puji dan syukur kita curahkan kepada Allah SWT Tuhan yang Maha Esa yang telah melimpahkan berkat dan rahmat-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul “Implementasi Metode *K-Nearest Neighbour* Dalam Peramalan Penjualan Mobil Bekas Di Kota Batam”. Skripsi ini merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa dalam penulisan skripsi ini masih banyak kekurangan dan tidak sempurna, oleh karena itu kritik dan saran dari pembaca sangat penulis harapkan.

Dalam penyusunan skripsi ini penulis mendapatkan bantuan dari berbagai pihak, maka dalam kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Batam
2. Dekan Fakultas Teknik dan Komputer Universitas Putera Batam
3. Bapak Andi Maslan, S.T., M.SI selaku Ketua Program Studi Teknik sekaligus pembimbing akademik pada program studi teknik Informatika Universitas Putera Batam selama perkuliahan.
4. Koko Handoko, S.Kom., M.Kom selaku pembimbing Proposal pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Dosen dan staff Universitas Putera Batam.

6. Kepada ayah yang sudah memberikan pesan dan saran agar penulis tetap melanjutkan kuliah sehingga sampai pada penghujung yaitu penyusunan skripsi ini.
7. Kepada ibu juga terima kasih yang selalu sabar dan memberikan motivasi saat penyusunan skripsi.
8. Saudara kandung kedua kakak dan abang yang selalu memberikan motivasi.
9. Tidak lupa adik tingkat penulis Rossalina sekaligus sebagai kekasih yang sudah memberikan semangat selalu.
10. Kepada teman-teman seperjuangan penulis yang sudah memberikan informasi maupun masukan dalam penyusunan skripsi ini.
11. Kepada semua orang yang menanyakan saya kapan wisuda.

Semoga Allah SWT membalas kebaikan dan selalu mencurahkan hidayah serta taufikNya, Amin

Batam, 27 January 2021

(M.Firmansyah Rafsyanzani s)



DAFTAR ISI

HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL	ii
HALAMAN PERNYATAAN	iii
HALAMAN PENGESAHAN.....	iv
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
DAFTAR RUMUS.....	xv
1.1 Latar Belakang Masalah	2
1.2 Identifikasi Masalah.....	4
1.3 Batasan Masalah.....	5
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian	6
1.6 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Knowledge Discovery In Database (KDD)	8
2.2 <i>Data Mining</i>	8
2.2.1 Pengelompokan <i>Data Mining</i>	10
2.2.2 Fungsi <i>Data Mining</i>	12
2.2.3 Kategori <i>Data Mining</i>	13
2.2.4 Manfaat <i>Data Mining</i>	14
2.2.5 Penerapan <i>Data Mining</i>	14
2.3 Metode <i>Data Mining</i>	15
2.3.1 Algoritma K-Nearest Neighbor	16
2.3.2 Konsep <i>K-Nearest Neighbour</i>	19
2.3.3 Fungsi <i>K-Nearest Neighbour</i>	19

2.4	<i>Software</i> Pendukung.....	20
2.5	Penelitian Terdahulu.....	20
2.6	Kerangka Pemikiran.....	24
BAB III METODE PENELITIAN.....		25
3.1	Desain Penelitian	26
3.2	Teknik Pengumpulan Data.....	28
3.2.1	Observasi	28
3.2.1	Wawancara.....	28
3.2.2	Studi Pustaka	28
3.3	Operasional Variabel	29
3.4	Metode Analisis Data	30
3.5	Lokasi Dan Jadwal Penelitian.....	30
3.5.1	Lokasi Penelitian	30
3.5.2	Jadwal Penelitian	31
BAB IV HASIL DAN PEMBAHASAN.....		32
4.1	Analisis Data.....	32
4.1.1	Data Penjualan.....	32
4.1.2	Praprocessing.....	35
4.2	Proses Pengujian.....	37
4.2.1	Model K-NN.....	37
4.2.2	K-Fold Cross-Validation.....	43
4.2.3	Pengujian Model	43
4.3	Pembahasan	55
BAB V SIMPULAN DAN SARAN		57
5.1	Simpulan	58
5.2	Saran.....	58
DAFTAR PUSTAKA.....		60
DAFTAR RIWAYAT HIDUP		64
SURAT IZIN PENELITIAN.....		66
LAMPIRAN.....		67

DAFTAR GAMBAR

Gambar 2.1 Tahapan KDD	8
Gambar 2.2 Kerangka Pemikiran	22
Gambar 3.1 Desain Penelitian	23
Gambar 4.1 Grafik Data Penjualan Mobil.....	32
Gambar 4.2 Lembar Kerja <i>Rapidminer</i>	35
Gambar 4.3 Import Dataset	35
Gambar 4.4 Pengaturan Tipe Data Pada Atribut	36
Gambar 4.5 Hasil Import Dataset	37
Gambar 4.7 Model <i>Cross Validation</i>	37
Gambar 4.8 Lembar Kerja Penempatan Model K-NN.....	38
Gambar 4.9 Model K-NN.....	39
Gambar 4.10 <i>Roc Optimization Precission</i>	43
Gambar 4.11 Grafik Perbandingan Prediksi Penjualan	46
Gambar 4.12 Grafik Pnejualn Tahun 2017.....	47
Gambar 4.13 Grafik Detail Penjualan Tahun 2017	47
Gambar 4.14 Grafik Penjualan Mobil Tahun 2018	48
Gambar 4.15 Grafik Detail Penjualan Tahun 2018.....	49
Gambar 4.16 Grafik Penjualan Mobil Tahun 2019	49
Gambar 4.17 Grafik Detail Penjualan Tahun 2019	50
Gambar 4.18 Grafik Penjualan Tahun 2020.....	51
Gambar 4.19 Grafik Detail Penjualan Tahun 2020	51

DAFTAR TABEL

Tabel 3.1 Jadwal Penelitian	28
Tabel 4.1 Data Penjualan.....	29
Tabel 4.2 Hasil Penjualan.....	31
Tabel 4.3 Hasil Pengolahan Penelitian.....	33
Tabel 4.4 K-Fold Cross-Validation	40
Tabel 4.5 Hasil Pengujian Model	40
Tabel 4.6 Confussion Matrix	42
Tabel 4.7 Data Confussion Matrix	43

DAFTAR RUMUS

Rumus 2.1 Perhitungan Jarak <i>Euclidean</i>	16
---	----



BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Masyarakat Indonesia sekarang ini sangat banyak dan memiliki kebutuhan yang berbeda-beda, mulai dari anak muda yang suka mobil gaul, orang tua yang ingin mobil untuk bepergian bersama keluarga, pedagang yang ingin mobilnya bisa mengangkut banyak barang, hingga pengusaha yang ingin mobil nyaman saat dikendarai. Berbagai macam kebutuhan masyarakat Indonesia tersebut ternyata bisa dipenuhi dengan berbagai variasi sesuai kebutuhan dan mudah didapatkan di Indonesia.

Mobil adalah alat kendaraan yang merupakan suatu kebutuhan oleh masyarakat kota yang berada pada level menengah ke atas. Seseorang banyak yang menginginkan atau memiliki mobil salah satunya masyarakat yang ada di kota Batam. Mobil-mobil yang ada di kota Batam berasal dari Singapura yang harganya bisa dijangkau oleh masyarakat yang ada di kota Batam. Mobil yang berasal dari Singapura memiliki harga yang murah dikarenakan bebas pajak penambahan nilai (PPN). Akan tetapi, mobil yang dibeli dari Singapura memiliki kekurangan tidak bisa dibawa keluar Batam. Untuk membawanya, pengendara wajib membayar pajak terlebih dahulu. Mobil-mobil yang berasal dari Singapura banyak ragam merek atau model. Jauh lebih murah dibandingkan dengan bekas yang ada di kota Batam. Hal ini disebabkan karena pemerintah Singapura memiliki kebijakan ialah umur maksimal mobil hanya 3 sampai dengan 5 tahun. Setelah itu, wajib diasingkan atau tidak boleh digunakan. Oleh

masyarakat batam dimanfaatkan menjadi barang seludupan dan batam juga sangat didukung oleh letaknya yang strategis. Sehingga para pemilik *showroom* banyak mendatangkan mobil dari Negara tersebut.

Penjualan mobil bekas pada salah satu *showroom* saat ini masih belum stabil dikarenakan belum begitu dilirik secara langsung di masyarakat kota Batam. Jenis mobil belum memenuhi keinginan serta merk mobil yang mempengaruhi ketersediaan spare part menjadi alasan minat masyarakat dalam menentukan kesukaan terhadap mobil yang akan dibeli, apabila mobil dalam jumlah besar tidak terjual maka penjualan tidak stabil maka mengakibatkan banyak mobil yang tidak laku sehingga terjadi penumpukan mobil dan tidak terawat dengan baik. Perawatan mobil tersebut juga menjadi terbengkalai. Sehingga untuk menyelesaikan permasalahan tersebut pimpinan *showroom* harus mampu melakukan prediksi peningkatan penjualan. Hasil dari prediksi atau peramalan bermanfaat untuk membantu investor dan manajemen dalam pengambilan keputusan investasi (Tiaratuti, 2014).

Dalam pengelolaan hasil penjualan mobil saat ini, seorang admin menggunakan perangkat komputer dan mencatat data transaksi dengan tujuan mendapat data sebagai laporan akhir mereka. Namun data tersebut belum dapat digunakan sebagai data yang valid dalam memprediksi hasil penjualan mobil di masa mendatang. Data transaksi pembelian mobil dengan cara konvensional masih diperuntukkan untuk mengetahui minat konsumen, perlu adanya cara untuk memprediksi hasil penjualan mobil untuk mengetahui minat masyarakat di kota

batam terhadap mobil yang diminati berdasarkan hasil transaksi sebelumnya, sehingga persediaan mobil dapat memenuhi permintaan konsumen.

Data yang akan diperoleh dalam proposal ini mencakup merk, harga, jenis mobil serta total penjualan per bulannya dalam tiga tahun terakhir. Data yang tersimpan dalam komputer dengan jumlah yang banyak ini digunakan sebagai *Data Mining*. *Data Mining* adalah aturan pencarian pola atau informasi menarik dalam suatu kumpulan data terpilih dengan menggunakan metode tertentu.

Data Mining yang bersumber dari data komputer dan transaksi penjualan kemudian diolah menggunakan algoritma *K-NN*. Algoritma *K-NN* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. (Abdul Rohman, 2012) *K-NN* termasuk algoritma *supervised learnin*, yang mana hasil *query instance* baru, diklasifikasikan berdasarkan mayoritas dari kategori pada *K-Nearest Neighbor(K-NN)*. Kelas yang paling banyak muncul, yang akan menjadi kelas klasifikasi (Abdul Rohman, 2012)).

1.2 Identifikasi Masalah

Berdasarkan uraian dari latar belakang dapat diidentifikasi masalah sebagai berikut :

1. Penjualan yang tidak stabil mengakibatkan banyak mobil yang akhirnya tidak laku dan tidak terawat.
2. Kurangnya minat konsumen membeli mobil bekas dikarenakan merk maupun sparepart dan jenis mobil yang kurang mencukupi.

3. Kapasitas mesin mobil yang tidak sesuai dengan pemakaian sehari-hari mengakibatkan konsumen secara tidak langsung enggan membeli.

1.3 Batasan Masalah

Untuk menghindari adanya penyimpangan dari judul dan tujuan yang sebenarnya, maka adapun batasan masalah yang ditemukan dalam penelitian ini sebagai berikut:

1. Lokasi Penelitian ini bertempat di *showroom Auto Grace*
2. Data yang digunakan dalam penelitian hanya empat tahun terakhir dari tahun 2017 sampai 2020.
3. Penelitian ini menggunakan *Algoritma K-Nearest Neighbor* dan *Rapid miner*

1.4 Rumusan Masalah

Berdasarkan permasalahan yang telah diungkapkan sebelumnya dapat dirumuskan masalah yang akan diselesaikan sebagai berikut:

1. Bagaimana memprediksi tingkat penjualan mobil bekas menggunakan algoritma *K-NN*?
2. Bagaimana *K-NN* melakukan prediksi penjualan mobil terlaris pada *showroom Auto Grace*?

1.5 Tujuan Penelitian

Berdasarkan Uraian dari rumusan masalah, maka tujuan dalam melakukan penelitian adalah:

1. Untuk memprediksi tingkat penjualan mobil menggunakan algoritma *K-NN*.
2. Untuk mengetahui penjualan mobil terlaris berdasarkan merk mobil menggunakan algoritma *K-NN*.

1.6 Manfaat Penelitian

Adapun manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Memberikan manfaat bagi perkembangan ilmu pengetahuan dan teknologi khususnya tentang *Data Mining* memprediksi hasil penjualan menggunakan algoritma *K-Nearest Neighbor*.
2. Penelitian ini dapat dijadikan referensi awal serta mengembangkan ilmu yang telah dipelajari selama perkuliahan jurusan teknik informatika
3. Dapat membantu dan mempermudah pihak *showroom* dalam perencanaan penyediaan unit.



BAB II

TINJAUAN PUSTAKA

2.1 Knowledge Discovery In Database (KDD)

KDD adalah proses nontrivial untuk mengidentifikasi pola atau hubungan yang valid, baru, berpotensi berguna, dan pada akhirnya dapat dipahami dalam kumpulan data untuk membuat keputusan penting (Fayyad, Piatetsky-shapiro, & Smyth, 1996)

Istilah Penemuan Pengetahuan dalam Basis Data, atau KDD singkatnya, mengacu pada proses luas dalam menemukan pengetahuan dalam data, dan menekankan aplikasi "tingkat tinggi" dari metode penambangan data tertentu. Ini menarik bagi para peneliti dalam pembelajaran mesin, pengenalan pola, database, statistik, kecerdasan buatan, akuisisi pengetahuan untuk sistem pakar, dan visualisasi data. Tujuan pemersatu dari proses KDD adalah untuk mengekstrak pengetahuan dari data dalam konteks database yang besar. Ini dilakukan dengan menggunakan metode data mining (algoritme) untuk mengekstrak (mengidentifikasi) apa yang dianggap pengetahuan, sesuai dengan spesifikasi ukuran dan ambang batas, menggunakan database bersama dengan preprocessing, subsampling, dan transformasi database yang diperlukan. (Choiriah, 2016).

2.2 Data Mining

Data Mining adalah proses menggunakan teknik atau metode tertentu untuk menemukan pola atau informasi yang menarik dalam data yang dipilih. Teknik, metode, atau algoritme dalam *Data Mining* sangat bervariasi. Pemilihan metode

atau algoritma yang tepat sebenarnya bergantung pada keseluruhan tujuan dan proses KDD, dan data yang digunakan sebagai model yang baik idealnya cukup sebagai data penelitian. Semakin banyak data, semakin sedikit kesalahan (error) dan semakin baik model yang digunakan sebagai patokan.

Menurut (Handoko, 2016). *Data Mining* merupakan sebuah aktivitas yang dilakukan dengan menggabungkan berbagai ilmu statistika, matematika, artificial intelligence dan machine learning dengan tujuan untuk melakukan ekstraksi dan identifikasi dari berbagai informasi yang ada sehingga menghasilkan suatu pengetahuan baru yang bermanfaat pada hasil kedepannya pada sebuah database yang besar.

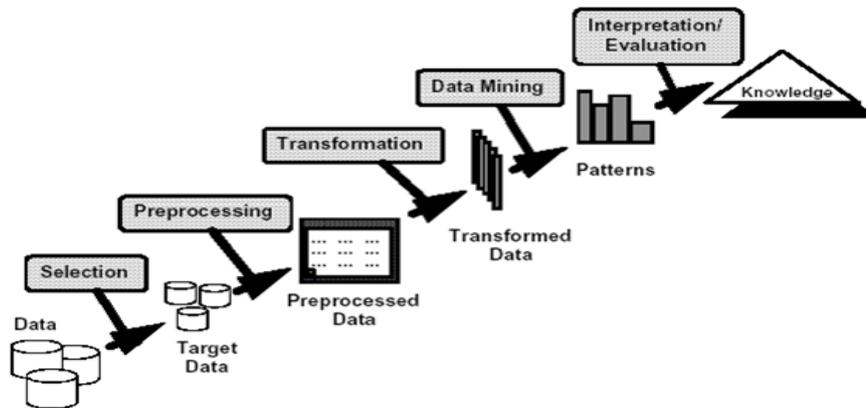
Data Mining adalah proses mendapatkan informasi yang berguna dari gudang database dalam bentuk pengetahuan. Penelitian ini menggunakan algoritma *Data Mining* dan *Rapidminer* untuk menganalisis data. Kemudian, *Data Mining* juga merupakan proses menemukan hubungan, pola, dan tren yang bermakna dengan memeriksa sejumlah besar data yang disimpan dalam memori menggunakan teknik pengenalan pola (seperti teknik statistik dan matematika).

(Thi Bi Dan et al., 2016)

Menurut (Budi Santoso, 2018) Kualitas data di ukur dari beberapa aspek. Aspek-aspek tersebut antara lain :

1. Akurasi: benar atau salah, akurat atau tidak
2. Kelengkapan: tidak tercatat, tidak tersedia
3. Konsistensi: beberapa sudah dimodifikasi yang lain belum
4. Terupdate (*timeliness*): apakah mengikuti waktu?

5. Bisa dipercaya: apakah bias dipercaya datanya benar?
6. Bisa diinterpretasikan: seberapa mudah datanya bias dipahami?



Gambar 2.1 Tahapan Knowledge Discovery In Database (KDD)

Sumber: *Data Mining Teori dan Aplikasi Rapidminer*

2.2.1 Pengelompokan *Data Mining*

Menurut Larose (2005) *Data Mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu sebagai berikut (Emha Taufiq Luthfi, 2009):

1. Deskripsi

Terkadang peneliti dan analis hanya ingin mencoba mencari cara untuk mendeskripsikan pola dan tren yang terdapat dalam data.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, hanya saja variabel target yang diestimasi lebih bersifat numerik daripada kategorikal. Buat model

menggunakan catatan lengkap yang memberikan nilai variabel target sebagai nilai prediksi. Selain itu, pada pemeriksaan selanjutnya, variabel target diestimasi berdasarkan nilai variabel prediktor. Contohnya adalah memperkirakan nilai kredit kumulatif mahasiswa pascasarjana dengan melihat nilai poin nilai mahasiswa yang mengikuti program sarjana..

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, hanya saja nilai hasil prediksi akan ada di masa yang akan datang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi juga dapat (jika sesuai) digunakan untuk prediksi.

4. Klasifikasi

Prediksi hampir sama dengan klasifikasi dan estimasi, hanya saja nilai hasil prediksi akan ada di masa yang akan datang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi juga dapat (jika sesuai) digunakan untuk prediksi.

5. Pengklusteran

Pengelompokan adalah pengelompokan catatan, pengamatan atau perhatian, dan membentuk suatu kelas dari objek yang serupa. Kluster adalah sekumpulan rekaman yang serupa satu sama lain tetapi tidak serupa dengan rekaman di kluster lain. Perbedaan antara clustering dan klasifikasi adalah tidak ada variabel target dalam clustering. Clustering tidak mencoba untuk mengklasifikasikan, memperkirakan atau memprediksi nilai variabel target. Akan tetapi, algoritma clustering berusaha untuk membagi seluruh data

menjadi kelompok yang serupa (seragam), dimana kesamaan record dalam satu kelompok akan memiliki nilai maksimal, dan kesamaan record dalam kelompok lain akan memiliki nilai minimum.

6. Asosiasi

Tugas korelasi dalam *Data Mining* adalah menemukan atribut yang muncul satu kali. Dalam dunia bisnis sering disebut juga analisis keranjang belanja. Analisis keranjang pasar adalah metode yang secara khusus menganalisis perilaku konsumen dari suatu kelompok / kelompok tertentu. Ketika kita tidak mengetahui pola spesifik yang kita cari, analisis keranjang pasar biasanya digunakan sebagai titik awal untuk menemukan pengetahuan tentang transaksi data. Teknologi analisis keranjang pasar merupakan teknologi yang menyesuaikan dengan ilmu *Data Mining*. Teknologi ini digunakan untuk merancang strategi penjualan dan pemasaran produk melalui proses pencarian untuk menemukan keterkaitan atau hubungan antara data item dalam database relasional. (Badrul, 2016).

2.2.2 Fungsi *Data Mining*

Menurut (Vulandari, 2017), fungsi - fungsi yang umum diterapkan dalam *Data Mining* (Lubis, 2016) yaitu :

1. Association, Ini adalah proses menemukan aturan asosiasi antara portofolio proyek.
2. Sequence, Proses menemukan aturan asosiasi untuk beberapa item sekaligus dan menerapkan beberapa periode.

3. Clustering, Merupakan proses pengelompokan beberapa data / objek ke dalam kelompok data sehingga setiap kelompok berisi data yang serupa.
4. *Classification*, Proses menemukan model atau fungsi yang mendeskripsikan atau membedakan konsep atau kategori data, dengan tujuan untuk dapat memperkirakan dari objek dengan label yang tidak diketahui..
5. Regression, adalah Proses memetakan data ke nilai prediksi.
6. Forecasting, adalah Proses memperkirakan nilai prediksi berdasarkan pola dalam kumpulan data.
7. Solution, adalah Proses menemukan akar penyebab masalah dan menyelesaikan masalah bisnis yang dihadapi atau setidaknya sebagai informasi untuk pengambilan keputusan.

2.2.3 Kategori Data Mining

Menurut (Vulandari, 2017), *Data Mining* dibagi menjadi dua kategori utama (Hand dan Kamber, 2006) yaitu:

1. Prediktif

Tujuan dari tugas prediksi adalah untuk memprediksi nilai atribut tertentu berdasarkan nilai atribut lain. Atribut yang diprediksi biasanya disebut variabel target atau variabel dependen, dan atribut yang digunakan untuk membuat prediksi disebut variabel penjelas atau variabel independen.

2. Deskriptif

Tujuan dari tugas deskriptif adalah untuk mendapatkan pola (korelasi, tren, kelompok, teori, dan anomali) yang meringkas hubungan yang sesuai

dalam kumpulan data. Tugas *Data Mining* deskriptif bersifat investigatif dan biasanya memerlukan teknik pasca-pemrosesan untuk memverifikasi dan menafsirkan hasilnya.

2.2.4 Manfaat *Data Mining*

Menurut (Vulandari, 2017), Pemanfaatan *Data Mining* dilihat dari dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan.

1. Dari perspektif bisnis, *Data Mining* dapat digunakan untuk mengatasi lonjakan volume data, dan teknologi komputasi dapat digunakan untuk menghasilkan informasi yang dibutuhkan. Ini merupakan aset yang dapat meningkatkan daya saing suatu organisasi.
2. Dari sudut pandang ilmiah, *Data Mining* dapat digunakan untuk menangkap, menganalisis, dan menyimpan data waktu nyata dan data yang sangat besar, misalnya :
 - a) *Remote sensor* yang ditempatkan pada suatu satelit.
 - b) *Teleskop* yang digunakan untuk memindai langit.
 - c) Simulasi saintifik yang membangkitkan data dalam ukuran *terabytes*.

2.2.5 Penerapan *Data Mining*

Menurut (Vulandari, 2017), Penerapan *Data Mining* dibagi menjadi tiga bagian yaitu sebagai berikut :

1. Analisa pasar dan kumpulan sumber data yang digunakan seperti transaksi kartu kredit, kartu anggota club tertentu, kupon diskon, keluhan pembeli, ditambah dengan studi dengan gaya hidup public. Beberapa solusi yang

dapat diselesaikan dengan *Data Mining* antara lain (menebak target pasar, melihat pola beli pemakai dari waktu ke waktu, cross market analysis, profil customer, identifikasi kebutuhan customer, menilai loyalitas customer, informasi summary)

2. Analisa perusahaan dan *management* resiko
 - a. Perencanaan keuangan dan evaluasi aset, *Data Mining* dapat membantu melakukan, analisis dan prediksi *cash flow* serta dapat melakukan *contingent claim analysis* untuk mengevaluasi aset.
 - b. Perencanaan sumber daya, dengan melihat ringkasan informasi serta pola pembelanjaan dan pemasukan dari masing-masing *resource*.
 - c. Persaingan, *Data Mining* dapat membantu untuk memonitor pesaing-pesaing dengan melihat *market direction* mereka.
3. Telekomunikasi, *Data Mining* melihat jutaan transaksi yang masuk dan melihat transaksi mana sajakah yang masih harus ditangani secara manual.

2.3 Metode *Data Mining*

Menurut (Yanto & Khoiriah, 2015). Analisis asosiasi atau penambangan aturan asosiasi adalah teknik penambangan data yang digunakan untuk menemukan aturan asosiasi di antara kombinasi ukuran kepentingan proyek yang dapat digunakan untuk penambangan data adalah:

1. *Support*, adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *item set* dari keseluruhan transaksi.
2. *Confidence*, adalah suatu ukuran yang menunjukkan hubungan antara dua

item secara conditional (berdasarkan suatu kondisi tertentu).

2.3.1 Algoritma K-Nearest Neighbor

Algoritma *K-Nearest Neighbour* terdekat (K-NN) merupakan metode pengklasifikasian objek berdasarkan data pembelajaran yang terdekat dengan objek tersebut. (Ilmu & Universitas, 2018)

Menurut (Mustakim & Oktaviani F, 2016) pembuktian sebuah Algoritma K-NN *Classification* yang mampu memodelkan sebuah kasus prediksi dengan akurasi yang tinggi. Algoritma *K-Nearest Neighbour* (K-NN) Ini adalah metode untuk mengelompokkan objek berdasarkan contoh pelatihan terdekat di ruang fitur. *K-Nearest Neighbour* adalah jenis pembelajaran berbasis contoh yang paling dasar atau pembelajaran malas, dan juga merupakan kelompok pembelajaran berbasis contoh. *K Nearest Neighbour* diselesaikan dengan menemukan kelompok objek terdekat (serupa) dalam data latih ke objek dalam data baru atau data uji. (Saxena et al., 2014).

Algoritme *K-Nearest Neighbour* sangat sederhana, berdasarkan Jarak terdekat dari sampel uji ke sampel pelatihan Tentukan tetangga terdekat K-nya. Setelah mengumpulkan K baru-baru ini Tetangga, kemudian mengambil mayoritas *K-Nearest Neighbour* (K-NN) sebagai Digunakan sebagai prediksi untuk sampel uji. *K-Nearest Neighbour* memiliki beberapa keunggulan Ini sulit untuk data pelatihan yang berisik dan efektif saat melatih data Besar. Selama fase pelatihan, algoritma ini hanya menyimpan vector Fungsi dan klasifikasi data sampel pelatihan. Pada tahap klasifikasi, fungsinya sama Perhitungan digunakan

untuk menguji data atau data yang klasifikasinya tidak diketahui. Jarak dari vector Hitungan baru dari semua vektor sampel pelatihan dan fragmen k yang paling dekat. Poin-poin yang memprediksi klasifikasi baru termasuk Poin-poin ini adalah yang paling rahasia. sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. (Sambanyu, 2016)

Menurut penelitian (Fatoni et al., 2018) algoritma *K-Nearest Neighbour* adalah suatu metode untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama berdasarkan bobot pencocokan dari beberapa fitur yang ada dengan kesamaan (similarity) Metode. Tujuan dari algoritma ini adalah untuk mengklasifikasikan objek baru berdasarkan atribut dan sampel pelatihan. Pengklasifikasi tidak menggunakan model apa pun untuk pencocokan, hanya berdasarkan memori.

K-Nearest Neighbor (K-NN) merupakan sebuah metode untuk melakukan klasifikasi terhadap obyek baru berdasarkan (K) tetangga terdekatnya (Abdul Rohman, 2012). K-NN termasuk algoritma *supervised learning*, yang mana hasil dari *query instance* baru, diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN. Kelas yang paling banyak muncul, yang akan menjadi kelas hasil klasifikasi (Abdul Rohman, 2012).

Pada algoritma K-NN terdapat 5 (lima) cara, untuk mencari tetangga terdekat (Sugesti et al, 2019) yaitu:

1. Jarak *Euclidean*
2. Jarak *Manhattan*

3. Jarak *Cosine*
4. Jarak *Correlation*
5. Jarak *Hamming*

Pada penelitian ini penulis hanya menggunakan jarak *Euclidean*, maka rumus perhitungan jarak dengan *Euclidean* seperti di bawah ini (Sayad 2010):

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Rumus 2.1 Perhitungan Jarak *Euclidean*

Nilai X_i merupakan nilai yang ada pada data *training*, sedangkan nilai Y_i merupakan nilai yang ada pada *testing*. Nilai K merupakan dimensi atribut.

Langkah-langkah untuk menghitung algoritma K -NN :

- 1) Menentukan nilai K
- 2) Menghitung kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap *training data* yang diberikan
- 3) Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil
- 4) Menggunakan label *class Y* (Klasifikasi *nearest neighbor*)
- 5) Dengan menggunakan kategori *nearest neighbor* yang paling mayoritas maka dapat diprediksikan nilai *query instance* yang telah dihitung

2.3.2 Konsep *K-Nearest Neighbour*

Algoritma *K-nearest Neighbour* (K-NN) adalah salah satu metode data mining yang paling populer (Wu & Kumar 2009). Metode K-NN murni termasuk dalam klasifikasi pelajar malas karena menunda proses pelatihan (atau bahkan tidak ada pelatihan sama sekali) hingga ada data pengujian yang ingin Anda ketahui label kategorinya, kemudian metode baru akan dijalankan. algoritme. Algoritma K-NN mengklasifikasikan berdasarkan kesamaan satu jenis data dengan jenis data lainnya. (Kurniawan & Saputra, 2019)

2.3.3 Fungsi *K-Nearest Neighbour*

Secara umum, *Data Mining* memiliki keterampilan untuk klasifikasi. Metode teknis ini mencakup dua teknik: pembelajaran yang diawasi daripada pembelajaran yang diawasi; pembelajaran yang diawasi bertujuan untuk menemukan pola-pola baru dalam data dengan menghubungkan pola data yang ada dan data baru; pada saat yang sama, dalam pembelajaran tanpa pengawasan, data tidak memiliki pola Dan gol. Belajar tanpa pengawasan untuk menemukan pola dalam data. Dalam studi pendaftaran ini, algoritma pembelajaran yang diawasi digunakan. Metode *K-Nearest Neighbour* didasarkan pada asumsi berikut: data memiliki kategori yang sama dengan data sekitarnya. Konsep ini disebut dengan konsep K-NN (Adi nugroho dan Sari, 2018). Menurut Gorunescu (Badu, 2016), *K-Nearest Neighbor* adalah metode klasifikasi objek baru berdasarkan tetangga terdekat (K)

2.4 *Software* Pendukung

Dalam *Data Mining*, kita dapat menggunakan berbagai *Software* dan beberapa contoh lainnya untuk mendukung *Data Mining* agar dapat mengambil keputusan berdasarkan tujuan dan proses dari data tersebut, oleh karena itu dalam penelitian ini peneliti akan menggunakan *Software* Rapidminer. Sesuai dengan yang diharapkan, ini adalah *Software Data Mining* gratis untuk keperluan akademik. Penelitian ini mengusulkan beberapa metode *Data Mining* dari bidang eksplorasi dan analisis data, pembelajaran statistik, pembelajaran mesin dan database. Rapidminer adalah perangkat lunak untuk memproses *Data Mining*. Tugas *Rapidminer* mencakup analisis teks, mengekstrak pola yang ada dari kumpulan data besar dan menggunakan metode statistik, kecerdasan buatan, dan database untuk menggabungkannya. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi bermutu tertinggi dari teks yang diolah.

2.5 Penelitian Terdahulu

Penelitian terdahulu merupakan acuan dalam sebuah penelitian, ada beberapa penelitian terdahulu yang digunakan dalam penelitian ini sebagai berikut.

1. Menurut Hwang dan Wen (1998:2026) dalam penelitiannya yang berjudul ***“FAST K-NN CLASSIFICATION AIGORITHM BASED ON PARTIAL DISTANCE SEARCH”*** Diketahui bahwa metode K-NN membutuhkan waktu kalkulasi yang cepat selama proses pencarian. Metode K-NN dianggap sangat efektif dalam aplikasi yang memerlukan tingkat kesalahan (di mana koneksi komputasi harus diwajibkan). Untuk aplikasi yang

membutuhkan akurasi tinggi dan distribusi abnormal, metode K-NN terbukti efektif dalam model statistik. Meskipun metode K-NN memiliki kelemahan pada jumlah vektor yang dibutuhkan untuk klasifikasi, metode K-NN dapat menghasilkan kompleksitas komputasi yang lebih tinggi untuk klasifikasi.

2. Menurut Wang, dkk (2013:1) dalam penelitiannya yang berjudul ***“EFFECIENT AND EFFECTIVE K-NN SEQUENCE SEARCH WITH APPROXIMATE N-GRAMS”*** menyatakan bahwa K-NN dapat mengatasi pemfilteran data pada database yang buruk. Selain itu, penelitian ini menunjukkan bahwa pencarian menggunakan K-NN lebih efisien dengan frekuensi antrian yang lebih sedikit. Penggunaan metode K-NN juga dapat memberikan efisiensi saat menghitung framework. Kesimpulan dari penelitian ini menunjukkan bahwa metode K-NN memberikan kinerja yang sangat baik dalam proses pencarian data.
3. Hasil Penelitian Prabowo (2014:47) ***“PEMBANGUNAN SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PENCARIAN LOKASI FITNESS CENTRE MENGGUNAKAN METODE K-NEAREST NEIGHBOUR”*** mengatakan bahwa dengan menggunakan metode K-NN dapat memberi skala berdasarkan kategori yang digunakan untuk metode perhitungan berdasarkan biaya anggota, jenis fasilitas, jumlah fitness center dan poin rating.
4. Swati dan Patankar (2014:157) dalam penelitiannya yang berjudul ***“EFFECTIVE PERSONALIZED MOBILE SEARCH USING K-NN”*** Tujuan dari penelitian ini adalah untuk memperbaiki klasifikasi dan

meminimalkan waktu pengerjaan. K-NN adalah algoritma yang digunakan untuk mengidentifikasi (lokasi dan konten) dan kemudian digabungkan untuk memberikan preferensi yang efektif kepada pengguna. Dalam penelitian ini dikatakan bahwa K-NN dapat meningkatkan klasifikasi lokasi dan konten berdasarkan konsep, sehingga kata kunci lokasi dapat dimaksimalkan dengan menambahkan semua kemungkinan kata kunci berbasis lokasi pada data training. K-NN dapat berada di server Memproses klasifikasi Di sisi sistem yang menggunakan algoritma K, NN menunjukkan fleksibilitas yang lebih besar dalam menyesuaikan preferensi pengguna. Algoritma K-NN yang digunakan dalam arsitektur ini jauh lebih cepat daripada algoritma SpyNB. Di sisi server, ini memiliki kecepatan pemrosesan yang lebih tinggi. Dalam penelitian ini dapat disimpulkan bahwa algoritma 11 K-NN digunakan untuk meningkatkan hasil klasifikasi dan memperoleh hasil yang lebih relevan. Kompleksitas algoritme K-NN lebih kecil daripada algoritme SpyBN, jadi jika data yang Anda miliki cukup besar, preferensi akan menemui beberapa kerumitan.

5. Komai et al (2015:942) dalam penelitian "***K NEAREST NEIGHBOUR SEARCH FOR LOCATION-DEPENDENT SENSOR DATA IN MANETS***" Dia mengatakan bahwa kueri K-NN digunakan untuk melakukan pencarian data sensor terbaru terkait lokasi penerbit kueri yang dapat digunakan pada LBS di lingkungan seluler. Adaptasi jaringan dan mobilitas node dapat memproses kueri lebih cepat. Metode K-NN dalam penelitian ini digunakan untuk mencari lokasi berdasarkan item data sensor

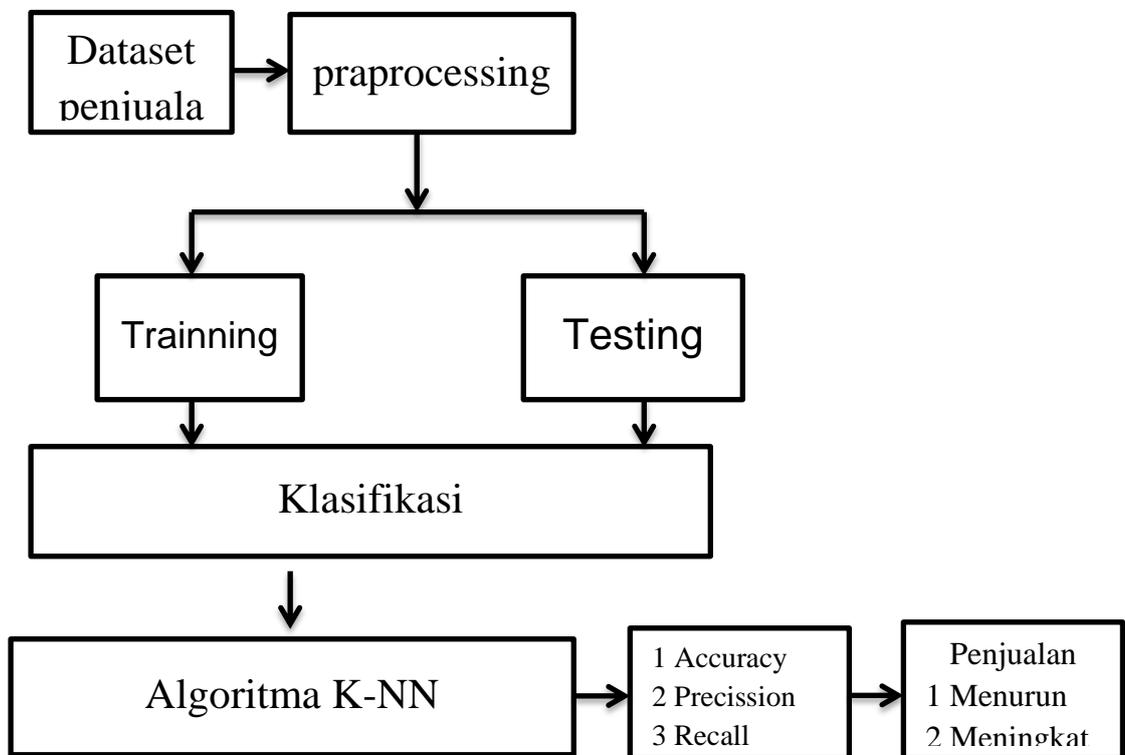
untuk mengurangi jumlah komunikasi dan menjaga keakuratan yang tinggi dari hasil query di MANET. Untuk mencapai rentang pencarian yang lebih kecil, item data disarangkan di dekat lokasi item dan node cache yang ditautkan. Yuka Komai et al. Menyimpulkan dalam penelitian ini bahwa metode K-NN memiliki overhead yang lebih rendah dan akurasi hasil query yang lebih tinggi.

6. Penelitian yang dilakukan oleh (Mega silfiani dan Suhartono., 2012) yang berjudul **“APLIKASI METODE ENSEMBEL PERAMALAN INFLASI DI INDONESIA”** metode single ANN menghasilkan RMSE yang kecil dibandingkan model lainnya pada pemodelan inflasi nasional dan tujuh kota di Jawa Timur. Sedangkan jika dilihat dari ensemble, metode ANN ensemble menghasilkan RMSE yang kecil dibandingkan model ensemble lainnya.
7. Penelitian yang dilakukan oleh (Nailah Fitriah., 2012) yang berjudul **“APLIKASI PREDIKSI HARGA SAHAM APPLE, IBM, DELL, DAN HP MENGGUNAKAN AIGORITMA KNEAREST NEIGHBOUR”** hasil penelitian ini Prediksi kenaikan atau penurunan harga saham Error rate (ratarata error terkecil) yang dihasilkan Algoritma KNearest Neighbors < 50%.
8. Penelitian yang dilakukan oleh (Ricky Imanuel Ndaumanu., 2014) yang berjudul **“ANALISIS PREDIKSI TINGKAT PENGUNDURAN DIRI MAHASISWA DENGAN METODE KNEAREST NEIGHBOUR”** Hasil

penelitian dengan uji coba secara manual dan menggunakan system didapatkan kesamaan hasil prediksi 79%.

2.6 Kerangka Pemikiran

Untuk memudahkan dalam proses pelaksanaan penelitian, maka berikut ini merupakan kerangka pemikiran dalam memprediksi penjualan di PT *Auto Grace* Kota Batam



Gambar 2.2 Kerangka Pemikiran



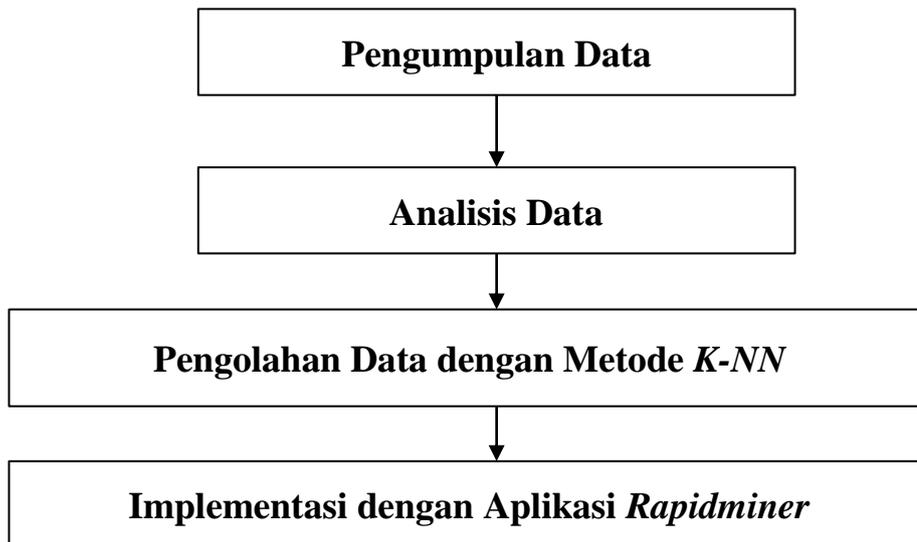
Universitas Putera Batam

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Dalam proses penelitian, terlebih dahulu perlu dibuat desain penelitian. Hal ini bertujuan untuk memberikan kemudahan penelitian yang lebih lanjut. Adapun desain dalam penelitian ini adalah sebagai berikut:



Gambar 3.1 Desain Penelitian

Sumber: Data Penelitian (2020)

Berdasarkan desain penelitian diatas yaitu suatu alur penelitian yang akan dilakukan dalam proses penelitian ini. Desain penelitian diatas menunjukkan proses dari awal kegiatan penelitian sampai akhir penelitian. Berikut ini adalah penjelasan dari desain penelitian diatas:

1. Pengumpulan Data

Pengumpulan data dalam penelitian ini menggunakan teknik observasi yaitu observasi langsung ke lokasi penelitian, selanjutnya melakukan wawancara dengan pihak accounting Auto Garage dan terakhir melakukan studi pustaka mengenai materi-materi *Data Mining* dan metode algoritma *K-NN* yang diperoleh dari buku-buku dan jurnal nasional maupun internasional.

2. Analisis Data

Analisa data yaitu proses dimana data yang diperoleh untuk menentukan operasional variabel yang dijadikan input dalam proses penelitian serta output yang dihasilkan sesuai dengan apa yang diinginkan.

3. Pengolahan Data dengan Metode *K-NN*

Data yang telah dianalisis selanjutnya akan diolah menggunakan teknik *Data Mining* menggunakan metode *Rapidminer*. Data yang diolah mencakup dua bagian yaitu merk mobil dan total penjualan mobil laku terjual. Untuk mendukung suatu keputusan pengolahan peneliti juga menambahkan tahun penjualan dan harga pada kolom data yang akan diteliti sebagai informasi tambahan dalam pengambilan keputusan, dan tidak termasuk bagian yang diolah.

4. Implementasi Dengan Aplikasi *Rapidminer*

Setelah semua data diolah maka data tersebut diimplementasikan kedalam aplikasi *Rapidminer*. Data yang diolah dengan aplikasi ini akan menjadi pertimbangan dalam perhitungan yang dilakukan secara manual.

3.2 Teknik Pengumpulan Data

Dari data yang akan diteliti, teknik pengumpulan data pada penelitian ini menggunakan tiga teknik yaitu teknik observasi, wawancara dan studi pustaka. Metode ini dilakukan secara berkala selama proses penelitian.

3.2.1 Observasi

Merupakan serangkaian proses biologis dan psikologis yang tersusun secara kompleks. Dua diantara terpenting adalah proses-proses pengamatan dan ingatan. Observasi dalam penelitian ini dilakukan secara langsung di Auto Garage untuk melakukan pengamatan terhadap data penjualan mobil bekas.

3.2.1 Wawancara

Merupakan salah satu instrumen yang digunakan untuk menggali data secara lisan. Hal ini haruslah dilakukan secara mendalam agar mendapatkan data yang valid dan detail. Wawancara dilakukan langsung pada pihak accounting Auto Garage.

3.2.2 Studi Pustaka

Dalam penelitian ini peneliti melakukan studi pustaka yang berhubungan dengan tema penelitian & metode penelitian yaitu jurnal dan buku yang berhubungan, untuk mencari informasi menyusun teori-teori yang berhubungan dengan pembahasan.

3.3 Operasional Variabel

Variabel penelitian pada dasarnya adalah segala sesuatu yang berbentuk apa saja yang ditetapkan oleh peneliti untuk dipelajari sehingga diperoleh informasi dan kesimpulannya (Sudaryono, 2015).

Operasional variabel dalam penelitian ini dapat dijelaskan dan diuraikan sebagai berikut:

1. Merk Mobil

Merk mobil merupakan jenis-jenis mobil yang paling banyak diminati dan dibeli oleh konsumen dan merupakan merk mobil sebagai pemilihan stok.

Merk mobil pada Auto Garage adalah Honda, Toyota, Mazda, Datsun, Daihatsu, Mitsubishi, Nissan, dan Suzuki.

2. Harga Jual

Disediakan untuk menunjang perbandingan harga yang ada dalam penjualan mobil terhadap harga mobil lainnya yang berbeda merek. Sehingga dapat menjadi pembanding pada merk lainnya. Harga yang bervariasi menyesuaikan tahun dan merk mobil yang ada.

3. Tahun Penjualan

Disediakan untuk menunjang perbandingan merk mobil dan harga jual dalam transaksi pemilihan mobil yang mana, tahun penjualan akan dijadikan salah satu kesimpulan bahwa tahun pembuatan mempengaruhi jual beli mobil dan minat konsumen di dalam memilihnya.

3.4 Metode Analisis Data

Dimulai dari data penjualan dari *Auto Garage* yang telah dianalisa oleh peneliti, data yang diperoleh yaitu data penjualan mobil di *Auto Garage* dari Januari 2017 sampai dengan Oktober 2020. Selanjutnya data tersebut dimasukkan ke dalam model K-NN, kemudian data-data akan dicari nilai maksimum support dari 1 itemset hingga akhir yaitu 10 itemset, setelah nilai maksimum support diperoleh maka langkah selanjutnya yaitu menentukan nilai *performance* sampai menghasilkan nilai final rulanya yang dilakukan secara manual menggunakan metode *K-NN*. Kemudian tahap terakhir yaitu melakukan pengujian dengan menggunakan aplikasi *rapidminer*, sebagai pembuktian bahwa hasil yang diperoleh secara manual sama dengan hasil implementasi dengan aplikasi.

Sedangkan menurut (Maslan et al., 2020), K-NN merupakan metode klasifikasi yang mampu mengklasifikasikan suatu atribut atau feature untuk mendapatkan tingkat akurasi yang paling optimal dan memiliki kecepatan analisis yang sangat cepat. Tool analisis yang digunakan Weka atau *Rapidminer*, dengan tujuan mempermudah seorang analis untuk memecahkan suatu permasalahan (*Problem Solving*).

3.5 Lokasi Dan Jadwal Penelitian

3.5.1 Lokasi Penelitian

Penelitian ini dilakukan di Auto Garage yang beralamat di Jl.Ahmad Yani, Taman Baloi, Kec.Batam Kota, Kota Batam.

3.5.2 Jadwal Penelitian

Penelitian ini dilakukan dalam waktu satu semester seperti yang telah dijadwalkan didalam kalender akademik, Tergantung pengajuan judul penelitian ini diterima atau tidaknya.

Tabel 3.1 Jadwal Penelitian

NO	NAMA KEGIATAN	BULAN KE				
		SEPT	OKT	NOV	DES	JAN
1	Survey Lokasi					
2	Mengidentifikasi Masalah					
3	Menentukan Tujuan Penelitian					
4	Mencari Referensi					
5	Menentukan Metode dan Mengumpulkan Data					
6	Menganalisis Data					
7	Membuat Laporan Penelitian					