

BAB 2

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Penelitian ini dilakukan tidak terlepas dari hasil penelitian-penelitian terdahulu yang pernah dilakukan sebagai bahan perbandingan dan kajian. Adapun hasil-hasil penelitian yang dijadikan perbandingan tidak terlepas dari topik penelitian yaitu mengenai peramalan.

Berdasarkan hasil penelitian yang pernah dilakukan Sholikah (2011) dimana melakukan penelitian mengenai peramalan tingkat penjualan sepeda motor Suzuki pada CV. Adi Mulia Motor di Rengat Indragiri Hulu. Peramalan menggunakan prosedur pembentukan model *time series Box-Jenkins*, diperoleh model yang paling sesuai yaitu AR(1). Hasilnya menunjukkan bahwa peramalan tingkat penjualan sepeda motor meningkat untuk periode 2011.

Destiarni (2018) menyajikan penelitian mengenai peramalan harga telur ayam ras pada hari besar keagamaan di Pasar Jawa Timur. Penelitian ini menemukan bahwa penerapan peramalan dapat digunakan untuk melihat pola fluktuasi harga telur ayam ras selama tiga tahun terakhir. Hasilnya kebijakan penetapan harga telur ayam ras yang ditetapkan secara nasional setiap tahunnya harus dikaji ulang.

Yuliana (2019) menyajikan penelitian mengenai peramalan penjualan pada PD. Sumber Jaya Aluminium. Penelitian ini menggunakan peramalan untuk menentukan metode peramalan yang tepat berdasarkan nilai *Mean Absolute*

Deviation (MAD), *Mean Square Error* (MER) dan *Mean Absolute Procentage Error* (MAPE). Diperoleh hasil penjualan rak piring menunjukkan kecenderungan naik dan metode yang dianggap terbaik terdapat pada metode *winter multiplikatif*.

Penelitian mengenai beberapa metode peramalan yang dilakukan oleh Hernadewita, dkk (2020) untuk mengetahui penjualan obat generik dan hasil peramalan penjualan obat generik periode selanjutnya pada Perusahaan Farmasi di Tangerang. Penelitian menggunakan metode peramalan *time series* dan teknik analisis data menggunakan *Microsoft Excel* dan *Minitab 17*. Hasilnya metode tren musiman sebagai metode terbaik untuk meramalkan penjualan obat generik.

Pada penelitian ini, Fadila (2021) menggunakan peramalan time dengan algoritma genetika untuk menentukan harga penjualan ayam dan hasil peramalan penjualan periode selanjutnya. Adapun perbandingan dari penelitian-penelitian terdahulu dapat dilihat pada tabel di bawah ini.

Tabel 2.1 Perbandingan dengan Penelitian-penelitian Terdahulu

	Parameter		
	Tujuan	Metode	Hasil
Sholikah (2011)	Peramalan tingkat penjualan sepeda motor Suzuki pada CV. Adi Mulia Motor	<i>Time series</i> <i>Box-Jenkins</i> .	Peramalan tingkat penjualan sepeda motor meningkat.
Destiarni (2018)	Peramalan harga telur ayam ras pada hari besar	Time series dengan metode ARIMA dan	Hasilnya kebijakan setiap

	keagamaan di Pasar Jawa Timur.	HW <i>exponential smoothing.</i>	tahunnya harus dikaji ulang.
Yuliana (2019)	Peramalan penjualan pada PD. Sumber Jaya Aluminium	<i>Time series</i> dengan metode MAD, MER dan MAPE	Hasil penjualan rak piring cenderung naik dan metode <i>winter multiplikatif</i> adalah yang terbaik
Hernadewita, dkk (2020)	Penjualan obat generik dan hasil peramalan penjualan obat generik periode selanjutnya	<i>Time series</i> dibantu dengan <i>Microsoft Excel</i> dan <i>Minitab 17</i>	Metode tren musiman sebagai metode terbaik
Fadila (2021)	Menentukan harga penjualan ayam berdasarkan hasil peramalan dari data PT. Integrasi Teknologi Unggas dan <i>website</i> www.pinsarindonesia.com	<i>Time series</i> dengan algoritma genetika	Diharapkan dapat menentukan harga penjualan ayam untuk periode selanjutnya

2. 2 Dasar Teori

2. 2. 1 Penjualan

Penjualan bertujuan untuk menyampaikan barang kebutuhan bagi yang memerlukan dengan imbalan uang menurut harga yang telah ditentukan. Di dalam praktik penjualan yang sehat, kegiatan penjualan barang terikat pada harga nyata karena dalam penentuan harga telah ditentukan oleh produksi dan konsumen (Widodo, 2005).

Adapun tujuan umum dari penjualan adalah sebagai berikut:

1. Mencapai volume atau jumlah penjualan tertentu
2. Mendapatkan laba tertentu
3. Menunjang pertumbuhan perusahaan (Widodo, 2005).

Faktor-faktor yang mempengaruhi tingkat penjualan antara lain:

1. Kualitas barang

Turunnya mutu barang dapat mempengaruhi tingkat penjualan, jika mutu suatu menurun, hal ini dapat menyebabkan konsumen merasa kecewa sehingga konsumen bisa berpaling kepada barang lain yang mutunya lebih baik.

2. Selera konsumen

Selera konsumen dapat berubah setiap saat, bilamana selera konsumen terhadap barang-barang yang kita perjualkan berubah maka tingkat penjualan akan menurun.

3. Servis terhadap konsumen

Servis terhadap konsumen merupakan faktor penting dalam usaha memperlancar penjualan dimana tingkat persaingan semakin tajam. Dengan adanya servis atau pelayanan yang baik terhadap para konsumen sehingga dapat meningkatkan tingkat penjualan.

4. Persaingan menurunkan harga jual

Potongan harga jual dapat dilakukan dengan tujuan agar penjualan dan keuntungan suatu perusahaan dapat ditingkatkan dari sebelumnya.

2. 2. 3 *Data Mining*

Menurut Hermawati (2013), *Data mining* adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengesktraksi pengetahuan (*knowledge*) secara otomatis.

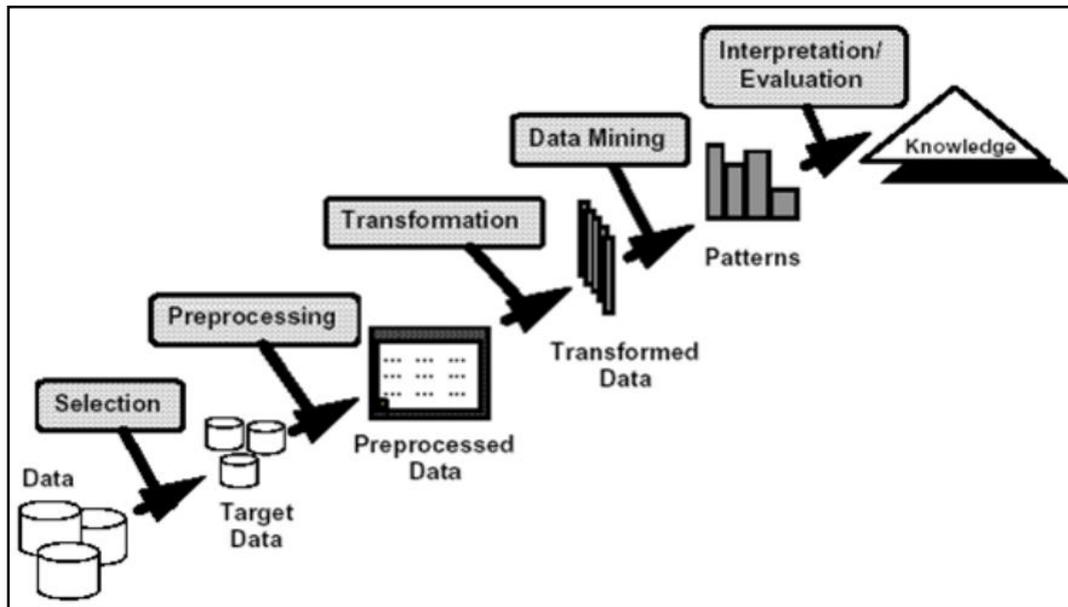
Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database besar* (Turban, dkk. 2005)

Berdasarkan definisi-definisi di atas tentang *Data mining* dapat disimpulkan bahwa *data mining* adalah sebuah proses pencarian secara otomatis untuk menemukan pola atau model dari suatu *database* yang besar.

2. 2. 4 *Knowledge Discovery in Database*

Knowledge Discovery in Database (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini

terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensi bermanfaat. *Data Mining* merupakan salah satu langkah dari serangkaian proses *iterative* KDD (Kusrini, dkk. 2009). Berikut tahapan proses KDD dapat dilihat pada gambar 2.1.



Gambar 2.1 Tahapan Proses KDD

Tahapan proses KDD terdiri dari :

1. *Data Selection*. Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target, atau memfokuskan pada subset *variable* (sampel data) dimana penemuan (*discovery*) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional
2. *Pre-Processing* dan *Cleaning Data*. *Pre-Processing* dan *Cleaning Data* dilakukan untuk membuang data yang tidak konsisten dan *noise*, duplikat data, memperbaiki kesalahan data, dan bisa diperkaya dengan data eksternal yang relevan.

3. *Transformation*. Proses ini *mentransformasikan* atau menggabungkan data ke dalam yang lebih tepat untuk melakukan proses *mining* dengan cara melakukan peringkasan (*agregasi*)
4. *Data Mining*. Proses *Data Mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode atau algoritma tertentu sesuai dengan tujuan dari proses KDD secara keseluruhan.
5. *Interpretation / Evaluasi*. Proses untuk menerjemahkan pola-pola yang dihasilkan dari *Data Mining*. Mengevaluasi (menguji) apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta atau hipotesa sebelumnya. Pengetahuan yang diperoleh dari pola-pola yang terbentuk dipresentasikan dalam bentuk visualisasi.

2. 2. 5 Runtun Waktu (*Time Series*)

Runtun waktu (*time series*) adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan perkembangan suatu kegiatan, misalkan sekumpulan data yang diambil per menit, per jam, per hari, per minggu, per bulan dan per tahun (Supranto, 2000). Contoh-contoh data yang diambil berdasarkan deretan waktu dalam kehidupan sehari-hari, diantaranya:

1. Tingkat penjualan sepeda motor.
2. Tingkat penjualan pakaian perhari di Matahari Plaza Pekanbaru
3. Tingkat produksi minyak perhari oleh PT. Cevron
4. Banyaknya penjualan HP Nokia per minggu untuk wilayah Pekanbaru

2. 2. 5 Model Linier *Time Series*

Secara umum model linier *time series* ada dua yaitu:

1. Model Linier *Time Series* yang Stasioner

Model ini dipakai untuk data yang stasioner tanpa melakukan proses *differencing* terhadap data. Model tersebut adalah (Hanke dkk, 2009):

a. Model *Autoregressive* Tingkat p atau AR(p)

Model AR adalah model linier yang paling dasar untuk proses yang stasioner, model ini dapat diartikan sebagai proses hasil regresi dengan dirinya sendiri. Secara umum model *autoregressive* dengan mengambil tingkatan p atau AR(p) didefinisikan sebagai berikut :

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (1)$$

b. Model *Moving Average* Tingkat q atau MA(q)

Model ini diartikan sebagai proses hasil regresi dengan errornya sendiri, bentuk umum dari proses *moving average* tingkat q atau MA(q) didefinisikan sebagai berikut :

$$Z_t = \phi_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2)$$

c. Model *Autoregressive Moving Average* tingkat p, q atau ARMA(p, q)

Model ini merupakan gabungan antara AR(p) dengan MA(q), sehingga dinyatakan sebagai ARMA(p, q) didefinisikan sebagai berikut :

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3)$$

2. Model *Linier Time Series Nonstasioner*

Model ini digunakan untuk data runtun waktu yang *nonstasioner*.

Adapun model tersebut adalah (Hanke dkk, 2009):

a. Model ARIMA (p, d, q)

Model ini merupakan suatu *time series nonstasioner* yang setelah diambil selisih (*differencing*) ke d menjadi stasioner, yang mempunyai model *autoregressive* tingkat p dan *moving average* tingkat q . Selanjutnya proses ARIMA yang tidak mempunyai bagian *moving average* ditulis sebagai ARIMA($p, d, 0$), dan ARIMA tanpa bagian *autoregressive* ditulis sebagai ARIMA($0, d, q$). Secara matematis model ARIMA(p, d, q), dengan $d = 1$ ditulis sebagai:

$$Z_t = \phi_0 + (1 + \phi_1)Z_{t-1} + (\phi_2 - \phi_1)Z_{t-2} + \dots + (\phi_p - \phi_{p-1})Z_{t-p} - \phi_p Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (4)$$

Keterangan untuk AR(p), MA(q), ARMA(p, q) dan ARIMA(p, d, q) adalah sebagai berikut :

Z_t	= data pada periode t , $t = 1, 2, 3, \dots, n$
ϕ_0	= konstanta
ϕ_i	= koefisien AR, $i = 1, 2, 3, \dots, p$
Z_{t-i}	= data pada periode $t - i$, $i = 1, 2, 3, \dots, p$
a_t	= <i>error</i> pada periode t , $t = 1, 2, 3, \dots, n$
θ_i	= koefisien MA, $i = 1, 2, 3, \dots, q$
a_{t-i}	= <i>error</i> pada periode $t - i$, $i = 1, 2, 3, \dots, q$
Z_{t-1}	= data pada periode $t - 1$, $i = 1, 2, 3, \dots, p$
a_t	= <i>error</i> pada periode t , $t = 1, 2, 3, \dots, n$

2. 2. 6 Web Scraping

Web Scraping adalah proses pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya berupa halaman-halaman Web dalam bahasa markup seperti HTML atau XHTML, dan menganalisis dokumen tersebut untuk diambil data tertentu dari halaman tersebut untuk digunakan bagi kepentingan lain (Turland, 2010). Tujuan utama dari web scraping adalah untuk mengumpulkan informasi atau data dari situs web dan menggunakannya untuk berbagai tujuan seperti analisis data, penelitian, pengumpulan informasi bisnis, atau membuat aplikasi yang memanfaatkan data tersebut. web scraping memiliki Langkah-langkah sebagai berikut:

1. Analisis Halaman Web:

Sebelum memulai web scraping, perlu untuk menganalisis struktur halaman web yang akan di-scrape. Identifikasi elemen-elemen HTML yang mengandung data yang ingin diekstraksi.

2. Inspect Element:

Gunakan fitur "Inspect Element" pada browser untuk melihat struktur HTML halaman web. Identifikasi elemen HTML, atribut, dan class yang berisi data yang diperlukan.

3. Pemilihan Metode Ekstraksi:

Tentukan metode ekstraksi yang sesuai dengan struktur HTML. Misalnya, menggunakan tag, class, atau XPath untuk mengidentifikasi elemen-elemen yang mengandung data.

4. Persiapkan URL:

Pastikan memiliki URL yang benar dan mengarah ke halaman web yang ingin di-scrape.

5. Request dan Response:

Lakukan HTTP request ke URL halaman web dan dapatkan responsnya. Analisis respons HTTP untuk memastikan bahwa halaman web memberikan data yang diinginkan.

6. Parsing HTML:

Proses parsing HTML respons untuk mengubahnya menjadi struktur data yang dapat diakses. Identifikasi elemen-elemen HTML yang mengandung data dan ekstrak informasi yang dibutuhkan.

7. Data Cleaning:

Setelah mendapatkan data, lakukan pembersihan data jika diperlukan. Ini mungkin melibatkan penghapusan karakter yang tidak diinginkan, penggabungan data, atau penyesuaian format.

8. Simpan Data:

Simpan data yang telah diekstraksi ke dalam format yang sesuai, seperti CSV, Excel, atau database, untuk memudahkan analisis dan penggunaan lebih lanjut.

2. 2. 7 RapidMiner

RapidMiner (YALE) adalah perangkat lunak *open source* untuk *knowledge discovery* dan *data mining*. *RapidMiner* memiliki kurang lebih 400 prosedur (operator) *data mining* termasuk operator untuk masukan, *output*, data *preprocessing* dan visualisasi (Sulianta,dkk 2010). Beberapa fitur dari *rapidminer*, antara lain:

1. Berlisensi gratis (*open source*)
2. Multiplatform karena diprogram dalam bahasa java.
3. Internal data berbasis xml sehingga memudahkan pertukaran data eksperimen.
4. Dilengkapi dengan scripting language untuk otomatisasi eksperimen.
5. Memiliki GUI (*Graphical User Interface*), command line mode (batch mode), dan Java API yang dapat dipanggil dari program lain.
6. Dapat dikembangkan dengan menambahkan plugin dan ekstension
7. Fasilitas plotting untuk visualisasi data multidimensi dan model.