

BAB 2

LANDASAN TEORI

2.1 *Forecasting (Prediksi)*

Prediksi adalah cara untuk memperkirakan bagaimana kondisi dan situasi yang sekarang akan mempengaruhi perkembangan di masa depan. Umumnya, orang menggunakan peramalan untuk melihat pola dari data yang sudah ada untuk menebak apa yang akan diperlukan di masa depan, dengan menggunakan metode tertentu, sehingga membantu memberikan gambaran permintaan yang sistematis. Proses ini bertujuan untuk memperkirakan banyak hal yang akan dibutuhkan ke depan, termasuk jumlah, kualitas, waktu, dan tempat yang diperlukan untuk memenuhi permintaan. Hal ini bertujuan untuk meminimalisir ketidakpastian kondisi di masa mendatang. *Forecasting* mirip dengan estimasi dan klasifikasi, yang membedakannya adalah bahwa prediksi berfokus pada memperkirakan suatu nilai di masa yang akan datang. Saat keadaan terbilang tepat, algoritma dan metode untuk estimasi dan klasifikasi juga dapat digunakan untuk melakukan prediksi (Johanes *et al.* 2024). Sementara itu, menurut Subagyo *Forecasting* atau pendugaan bertujuan untuk mendapatkan ramalan yang akurat untuk masa depan dengan melihat nilai kesalahan dari ramalan tersebut, yang dihitung dengan menggunakan ukuran seperti MAD (Mean Absolute Deviation), MSD (Mean Squared Deviation), dan MAPE (Mean Absolute Percentage Error).

2.2 *Data Mining*

Data Mining merupakan perhitungan matematis dan statistik yang bertujuan untuk mengidentifikasi informasi dari kumpulan data berukuran besar. Ini melibatkan teknik seperti penggalian pola, tren, atau pengetahuan baru melalui proses seperti klasifikasi, asosiasi, regresi dan *Time Series*. *Data Mining* digunakan untuk memproyeksikan tren dan karakteristik bisnis dengan memproses informasi dari big data, membantu menemukan semua pola yang sebelumnya tidak diketahui dengan menganalisis data, dan juga berperan dalam pengambilan sebuah keputusan dan strategi (Johanes Surya Putra *et al.* 2024). Beberapa tahapan *Data Mining* yang

meliputi proses dalam penambangan data seperti seperti membersihkan data, menggabungkan data, memilih data, mengubah data, dan menganalisis pola untuk mendapatkan informasi dari data itu. Secara umum, Data Mining terbagi menjadi dua jenis besar, yaitu yang bersifat prediktif dan yang bersifat deskriptif. *Prediktif* adalah proses untuk menemukan pola tertentu berdasarkan data historis dengan memanfaatkan variabel tertentu yang nantinya pola tersebut dapat diketahui dari berbagai variabel yang ada pada data, contoh teknik dalam prediktif adalah klasifikasi, regresi dan juga *Time Series*. Sedangkan *Deskriptif* adalah proses menemukan pola atau karakteristik penting dari data yang ada tanpa memprediksi masa depan yang berfungsi untuk memahami lebih jauh tentang data yang diamati, contoh teknik dalam deskriptif adalah clustering, association, sequential pattern mining. Sederhananya *Data Mining* bisa dikatakan sebagai proses penyaring atau “menambang” pengetahuan dari sejumlah data yang besar (Sri & nurwati, 2022).

2.3 *Time Series*

Time Series (deret waktu) adalah sekumpulan data dikelompokkan dalam urutan berdasarkan waktu dengan jeda yang tetap, seperti setiap hari, setiap bulan, atau setiap tahun. Data ini digunakan untuk menganalisis pola tren, atau perilaku suatu fenomena dari waktu ke waktu serta membuat prediksi (*Forecasting*) dimasa mendatang. Analisis kebutuhan sebelumnya digunakan untuk memilih metode, dan data yang akan digunakan adalah data penjualan berurutan berdasarkan waktu.

Least Square adalah salah satu dari banyak metode turunan metode *Time Series*. Ada banyak hal yang mempengaruhi data seri waktu, menurut M. Fariz Aminulloh (2019). Dalam data ekonomi, fluktuasi dari waktu ke waktu biasanya disebabkan oleh faktor tren (faktor tren), fluktuasi siklis (fluktuasi siklis), variasi musiman (variasi musiman), dan pengaruh acak.

2.4 *Metode Least Square*

Metode *Least Square*, juga dikenal sebagai kuadrat terkecil adalah cara untuk memprediksi apa yang akan terjadi di masa depan dengan menggunakan garis lurus yang sesuai dengan data sebelumnya. Metode ini membantu kita Untuk

memperkirakan apa yang akan terjadi di masa depan, teknik ini membantu kita memahami data yang telah terjadi (Ridwan, Faisol, & Wahyuni, 2020). Metode ini terutama digunakan untuk menganalisis data deret waktu, yaitu data yang terjadi secara berurutan sepanjang waktu. Untuk menggunakan metode ini, kita membutuhkan data lama untuk bisa menebak kondisi yang akan datang. Ada dua cara untuk menggunakan metode ini, yaitu untuk jumlah data genap dan data ganjil, cara ini berbeda tergantung pada bagaimana kita menentukan nilai X , yang menunjukkan periode waktu. Metode ini menghasilkan garis lurus yang mengurangi jumlah kuadrat dari perbedaan antara data nyata dengan data yang diprediksi (Dwi Ika Pebri *et al.* 2021). Garis kuadrat terkecil dapat digambarkan dengan titik di sepanjang sumbu Y di mana garis tersebut bersilangan. Bentuk persamaan dari metode Least Square adalah sebagai berikut:

$$\hat{Y} = a + b(X) \quad (2.1)$$

Keterangan :

- \hat{Y} : Hasil Prediksi
- X : Periode (waktu)
- a : Bilangan konstan
- b : Koefisien kecondongan garis tren

Untuk menghitung data, variabel waktu (X) yang berfungsi sebagai bobotnya harus menerima nilai tertentu. Variabel waktu yang berada di tengah diberi nilai 0, seperti yang ditunjukkan oleh contoh berikut:

1. Pada kasus data ganjil jarak antara dua waktu diberi nilai satu satuan dari titik tengah. Contoh untuk 9 data adalah (... , -3, -2, -1, 0, 1, 2 , 3,).
2. Jika jumlah data genap jarak antara dua waktu diberi nilai dua satuan. Contoh untuk 8 data maka skor nilai nya adalah (... -5, -3, -1, 1, 3, 5,).

Untuk mendapatkan nilai Y , koefisien a dan b harus dicari terlebih dahulu dengan menggunakan persamaan sebagai berikut :

$$a = \frac{\sum Y}{n} \quad (2.2)$$

$$b = \frac{\sum XY}{\sum X^2} \quad (2.3)$$

Keterangan :

- Y : Data berkala
- X : Periode (waktu)
- a : Bilangan konstan
- b : Koefisien kecondongan garis tren
- n : Jumlah data

Setelah didapat nilai dari koefisien a dan b maka nilai tersebut digunakan dalam persamaan 2.1 untuk memprediksi data yang akan datang.

2.5 Perhitungan Error

Metode perhitungan *Error* dalam analisis data sangat penting untuk mengukur akurasi suatu model prediksi atau peramalan. *Mean Absolute Percentage Error (MAPE)* dan *Root Mean Square Error (RMSE)* adalah dua metode evaluasi yang sering digunakan dalam analisis prediksi atau peramalan. Meskipun keduanya mengukur seberapa jauh hasil prediksi dari nilai aktual, mereka memiliki perbedaan mendasar dalam cara menghitung dan interpretasinya.

1. Mean Absolute Percentage Error (MAPE)

MAPE adalah cara untuk menghitung perbedaan antara data yang sebenarnya dan data yang diprediksi. Perbedaan ini diubah menjadi angka positif, dan lalu kita menghitung rata-rata kesalahan dalam bentuk persentase dari nilai yang sebenarnya sehingga hasilnya bisa dinyatakan dalam persen. Dari hasil persentase ini, kita kemudian mendapatkan nilai rata-ratanya. *MAPE* lebih banyak digunakan untuk mengukur tren (Nurkahfi *et al.* 2020). Rumus *MAPE* adalah sebagai berikut :

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{Y_t} \times 100 \quad (2.4)$$

Keterangan :

Y_t = Data aktual pada periode ke t

\hat{Y} = Nilai ramalan pada periode ke t

n = Jumlah data

Semakin kecil angka persentase *MAPE*, semakin baik kinerja model ramalan yang dipakai. *MAPE* memiliki nilai-nilai tertentu yang bisa digunakan untuk menilai seberapa efektif suatu model ramalan.

2. Root Mean Square Error (RMSE)

RMSE adalah metrik evaluasi yang digunakan untuk mengukur seberapa baik keakuratan hasil perhitungan yang telah dilakukan dengan perbandingan tingkat kesalahan. *RMSE* mengukur seberapa besar deviasi antara nilai-nilai yang diprediksi oleh model dan nilai-nilai yang ditemukan dalam data yang sebenarnya. Nilai aktual dan nilai prediksi dapat dihitung secara matematis dengan menggunakan *RMSE* untuk menghitung akar dari rata-rata kuadrat. Adapaun rumus *RMSE* sebagai berikut :

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y})^2}{n}} \quad (2.5)$$

Keterangan :

Y_t = Nilai data aktual

\hat{Y} = Nilai hasil peramalan

n = Jumlah data

RMSE adalah salah satu cara umum untuk menilai hasil dalam analisis regresi. Ini membantu kita mengukur seberapa baik model bekerja dan juga membandingkan seberapa bagus berbagai model regresi yang berbeda. Semakin kecil nilai *RMSE*, semakin baik model regresi dalam melakukan prediksi (Agus Julkarnaen *et. al.* 2024).

2.6 Penelitian Terdahulu

Berikut adalah beberapa referensi pembelajaran dari beberapa contoh kasus yang memiliki masalah yang hampir sama dengan yang dihadapi, antara lain:

1. (Danendra Benny *et al.* 2025) dengan judul “Penggunaan Metode Economic Order Quantity dan *Least Square* Pada Aplikasi Prediksi Peramalan Obat Berbasis Web”.

Fokus penelitian ini adalah mengatasi masalah manajemen persediaan obat di Klinik Pratama yang masih dilakukan secara manual. Masalah-masalah ini sering menyebabkan kelebihan atau kekurangan stok, pola permintaan yang tidak menentu, dan ketidakakuratan dalam jumlah pembelian. Metode *Least Square* dapat digunakan untuk membuat estimasi permintaan obat yang sangat mirip dengan data aktual. Ini menunjukkan bahwa sistem dapat diandalkan untuk membuat keputusan tentang persediaan obat di klinik (Danendra Benny *et al.*, 2025).

2. (Muh. Naufal *et al.* 2022) dengan judul “Rancang Bangun Aplikasi Peramalan Penjualan Berbasis Website Menggunakan Metode *Time Series*”.

Penelitian ini membahas tentang peramalan penjualan menggunakan metode *Time Series*. Dari hasil peramalan dengan metode *Least Square* terlihat bahwa penjualan cenderung meningkat dari 26 unit ke 28 unit, berbeda dengan metode *single exponential smoothing* yang hasilnya masih bervariasi dan naik turun. Oleh karena itu, untuk data fluktuatif lebih dianjurkan menggunakan metode *single exponential smoothing* dibandingkan metode *Least Square* (Muh. Naufal *et al.* 2022).

3. (Indah Sari *et al.* 2022) dengan judul “Peramalan Prediksi Penjualan Garam Pada CV. Saltindo Megajaya Dengan Metode *Least Square*”.

Penelitian tersebut membahas tentang peramalan penjualan garam pada perusahaan. Hasil penelitian tersebut menyebutkan bahwa didapatkan hasil prediksi yang secara cepat dan akurat berdasarkan data-data bulan sebelumnya dan dari pengujian yang dilakukan dengan menggunakan Metode *Least Square* didapat perhitungan *Error (MAPE)* sebesar 6% dengan hasil prediksi untuk bulan Juni 2022 sebesar 23.810 Kg (Indah Sari *et al.* 2022).

4. (Qiqi Ari Fuzaki 2022) dengan judul “Perbandingan Metode Double Exponential Smoothing Dan *Least Square* Untuk Prediksi Penjualan Kue”.

Penelitian tersebut membahas tentang hasil perbandingan metode manakah yang lebih efektif memiliki nilai yang paling baik dalam memprediksi penjualan kue. Berdasarkan tahapan pengujian bahwa metode Double Exponential Smoothing dan *Least Square* yang diimplementasikan didapatkan hasil perbandingan rata-rata nilai *Error* yang berbeda. Pada metode *Least Square* memiliki nilai terkecil dengan hasil akurasi yang lebih akurat sebesar 15,49%, sedangkan pada metode Double Exponential Smoothing sebesar 17,48% (Qiqi Ari Fuzaki 2022).

5. (Dimas Fahmi *et al.* 2023) dengan judul “Faktor-faktor yang Mempengaruhi Penjualan Obat Bebas dan Bebas Terbatas di Apotek Hafshawaty”.

Penelitian ini menganalisis tren penjualan obat bebas dan terbatas di Apotek Hafshawaty selama Desember 2023. Hasil studi menunjukkan bahwa obat pencernaan dan pernapasan mendominasi penjualan, mengindikasikan prevalensi masalah kesehatan tersebut di masyarakat sekitar. Kesimpulannya, artikel ini menyoroti pentingnya akses mudah, edukasi apoteker, dan ketersediaan obat dalam mempengaruhi pola swamedikasi masyarakat, serta perlunya edukasi berkelanjutan untuk penggunaan obat yang rasional. (Dimas Fahmi *et al.* 2023).

6. (Munich Heindari *et al.* 2021) dengan judul “Perancangan Sistem Perkiraan Metode *Least Square* Pada Penjualan Produk Keramik”.

Penelitian tersebut membahas tentang perancangan sistem perkiraan metode *Least Square* untuk prediksi penjualan produk keramik. Aplikasi yang dirancang menggunakan metode *Least Square* ini dapat memperkirakan pembelian bahan baku untuk membuat produk yang dijual sesuai dengan kondisi penjualan barang, sehingga memberikan informasi berhubungan dengan target jumlah produksi yang lebih mendekati angka penjualan bulan yang berjalan. (Munich Heindari *et al.* 2021).

7. (Riki Maulana *et al.* 2021) dengan judul “Implementasi *Data Mining* Menggunakan Metode *Least Square* Untuk Memprediksi Penjualan Lampu Led Pada Pt. Sumber Dinamika Solusitama”.

Penelitian ini membahas penerapan teknik data mining untuk memprediksi penjualan lampu LED di sebuah perusahaan. Metode yang digunakan adalah *Least Square*, yang menghasilkan prediksi yang lebih tepat. Prediksi ini membantu perusahaan dalam pengambilan keputusan pembelian dan pengadaan persediaan. Untuk mengetahui tingkat akurasi prediksi, penulis menggunakan perhitungan *Mean Absolute Percentage Error*, yang menghasilkan nilai 8,0744%. Berdasarkan nilai tersebut, dapat disimpulkan bahwa model peramalan yang digunakan memiliki kemampuan yang sangat baik (Riki Maulana *et al.* 2021).

8. (Dayu *et al.* 2022) dengan judul “Analisis Perbandingan Metode *Least Square* Dan Parabolik Untuk Perhitungan *Forecasting* Penjualan Minuman Pada Kedai Rumah Celoteh”.

Penelitian tersebut membahas tentang analisis perbandingan metode *Time Series* dan metode *parabolik* untuk perhitungan prediksi penjualan minuman pada kedai. Dari pengujian menghasilkan metode *Least Square* nilai secara keseluruhan adalah 28,48% dengan kesalahan tiap data adalah 3,16%, dan tingkat kesalahan SKP adalah Rp 511,141, sedangkan *Parabolik MAPE* secara keseluruhan adalah 27,45% dengan kesalahan tiap data adalah 3,05%, dan tingkat kesalahan SKP adalah Rp 503,441.23. Hasil tersebut didapat nilai kesalahan yang relatif kecil, sehingga didapat kesimpulan bahwa menggunakan kedua metode bisa menjadi metode yang baik untuk prediksi (Dayu *et al.* 2022).

9. (Hani Hamidah *et al.* 2024) dengan judul “Prediksi Pemenuhan Kebutuhan Darah Dengan Metode *Least Square* Di Utd Pmi Kabupaten Cianjur Jawa Barat”.

Penelitian tersebut bertujuan untuk mengetahui apakah metode *Least Square* bisa digunakan dan kemampuan prediksi metode *Least Square* dalam memprediksi pemenuhan kebutuhan darah di UTD PMI. Hasil

menunjukkan nilai kesalahan prediksi terendah yaitu 0,6% pada prediksi PRC bulan Oktober dan nilai kesalahan prediksi tertinggi yaitu 3,6% pada prediksi WB bulan September. Dari hasil tersebut dapat disimpulkan bahwa metode *Least Square* dapat digunakan untuk prediksi pemenuhan darah di UTD PMI dengan tingkat kemampuan prediksi yang sangat baik karena kesalahan <10%. (Hani Hamidah *et al.* 2024).

10. (Ni Made Ayu *et al.* 2024) dengan judul “Perbandingan Metode *Least Square* Dengan Metode Moment Dalam Menyusun Anggaran Penjualan dan Implikasinya Terhadap Perencanaan Laba di UD Timbul Variasi”.

Penelitian ini bertujuan untuk membandingkan metode peramalan kuadrat terkecil dengan metode momen serta menghitung SKP yang terjadi. Hasil menunjukkan bahwa peramalan penjualan tahun 2023 menggunakan metode momen lebih sesuai dengan penjualan aktual, yaitu sebesar 2.873 dengan kesalahan standar peramalan sebesar 27,45. Selisih SKP sebesar 0,1 dianggap tidak signifikan. Oleh karena itu, metode momen cocok digunakan dalam perusahaan. Kesimpulan yang diperoleh adalah metode naive bayes bisa menjadi pilihan yang tepat untuk digunakan dalam proses seleksi karyawan untuk bagian pemasaran (Ni Made Ayu *et al.* 2024).



GRESIK