BAB II
KAJIAN PUSTAKA

2.1 Tinjauan Pustaka

Penelitian dengan topik peramalan telah dilakukan oleh banyak peneliti pada saat ini. Peneliti telah melakukan penelitian di berbagai bidang dan menggunakan berbagai macam metode prediksi (prediksi rentet waktu). Metode prediksi yang digunakan misalnya ARIMA, *Eksponensial Smoothing*, *Moving Average* dan metode-metode *Machine Learning*, khususnya *Neural Network*. Metode-metode ini juga telah sukses digunakan untuk prediksi saham.

Nurdin (2022) telah melakukan penelitian dengan menghasilkan sistem aplikasi yang mampu mengolah data historis indeks saham LQ45 untuk memprediksi pergerakan harga di masa depan. Metode yang digunakan adalah *Triple Exponential Smoothing* yang merupakan metode peramalan berbasis analisis statistik. Hasil penelitian menunjukkan bahwa *Triple Exponential Smoothing* efektif diterapkan untuk prediksi harga saham dengan tingkat *mean margin error* masing-masing: *Open* 0,10681%, *High* −1,1156%, *Low* 1,4616%, dan *Close* −0,2504%. Dengan demikian metode *triple Exponential Smoothing* dapat digunakan untuk memprediksi pergerakan harga indeks LQ45 dengan kesalahan yang relatif kecil.

Saputra & Suharsono (2025) telah berhasil menggunakan model ANN untuk memprediksi harga saham PT Bank Central Asia Tbk. Penelitian ini menegaskan bahwa ANN dapat mengenali pola pergerakan harga saham perusahaan perbankan dengan baik dan memberikan hasil prediksi yang akurat. Widhi Aryanti & Nur Azizah Komara Rifai (2023)juga telah menggunakan model ANN dengan dalam prediksi harga saham. Hasil penelitian juga membuktikan bahwa model ANN dapat menyesuaikan bobot jaringan secara dengan optimal selama proses pelatihan, dan dapat meningkatkan kinerja model prediksi dibandingkan metode linier. Penelitian -penelitian ini menunjukkan bahwa ANN efektif dalam memodelkan prediksi harga saham sektor perbankan yang cenderung dinamis.

Bao dkk (2025) melakukan tinjauan komprehensif tentang *Neural Network* berbasis data dalam prediksi saham. Metode yang dilakukan dalam penelitian ini meliputi tinjauan komprehensif literatur yang ada tentang model-model *Neural Network* berbasis data dalam prediksi saham, pencapaian model-model dalam prediksi pasar saham, dan metrik evaluasi yang umum digunakan di bidang peramalan saham. Metode yang digunakan dalam penelitian ini meliputi tinjauan komprehensif terhadap 63 artikel jurnal dan konferensi dari tahun 2015 hingga 2023. Hasil dalam penelitian ini menunjukkan bahwa *Neural Network* berbasis data efektif dalam prediksi harga saham.

Jin (2025) melakukan penelitian dengan mengeksplorasi hubungan pemetaan nonlinier antara data saham historis dan harga penutupan (close) hari berikutnya menggunakan model *Backpropagation Neural Network* (BPNN). Temuan dalam penelitian ini adalah bahwa model BPNN menunjukkan kemampuan peramalan yang kuat di pasar yang stabil, dan memperpanjang rentang waktu dapat meningkatkan akurasi prediksi. Penelitian ini mendapatkan kinerja model BPNN dengan R^2 mencapai hingga 0,96.

Rahim dkk. (2025) telah menggunakan teknik machine learning dengan model Jaringan Saraf Buatan (*Artificial Neural Network*s/ANN) untuk memprediksi harga saham, khususnya memprediksi harga saham Apple Inc. (AAPL) dan *Microsoft Corp*. secara akurat. Dari penelitian ini ditemukan bahwa model ANN dapat memprediksi harga saham secara akurat dan andal, menangkap tren yang mendasari pergerakan harga saham. Dari penelitian ini dahasilkan bahwa model ANN dengan satu lapisan tersembunyi dapat memprediksi harga saham secara akurat, dengan nilai RMSE yang rendah sebesar $1,474 +/- $0,114 untuk harga saham AAPL.

Singh dkk (2024) melakukan penelitian dengan mengevaluasi kinerja berbagai model regresi, seperti Regresi Linier, Regresi Pohon Keputusan, Regresi *K-Nearest Neighbor*, Regresi *Multilayer Perceptron*, Regresi *Support Vector*, dan lain-lain. Hasil penelitian menunjukkan bahwa penerapan pemilihan fitur dengan metode *Forward Selection* secara signifikan meningkatkan kinerja model, dengan peningkatan sebesar 56,47% pada Regresi *Multilayer Perceptron*. Dengan kata lain

MLPSF (Integrasi MLP dan *Forward Selection*) mengungguli model individual MLP sebesar 56,47%.

Metode *tuning hyperparameter Grid Search* telah berhasil meningkatkan kinerja model *machine learning* untuk prediksi saham. Peneliti (Salsabilla dkk., 2024) telah prediksi harga penutupan saham (*close*) PT Indofarma (INAF) selama 30 hari ke depan menggunakan metode *Support Vector Regression* (SVR). Setelah menggunakan *tuning hyperparameter Grid Search*, model SVR dengan kernel RBF dan penyetelan hyperparameter dapat memprediksi harga penutupan harian saham INAF secara akurat, dengan kinerja nilai *Mean Absolute Percentage Error* (MAPE) yang rendah yaitu 1,537%.

2.2 Landasan Teori

2.2.1 Pengertian Saham

Saham didefinisikan sebagai surat berharga yang mewakili bukti kepemilikan atas sebagian ekuitas dari suatu perusahaan perseroan terbatas (Bodie dkk., 2021). Pemegang saham (*investor*) memiliki klaim terhadap aset dan pendapatan perusahaan sesuai dengan proporsi kepemilikan sahamnya. Keuntungan dalam investasi saham berasal dari dua sumber utama, yaitu *capital gain* (apresiasi harga) dan *dividen* (pembagian laba) (Berk & DeMarzo, 2023).

Saham dapat dikelompokkan berdasarkan beberapa kriteria. Berdasarkan hak klaim, terdapat saham biasa (*common stock*) yang memberikan hak suara dan dividen variabel, serta saham preferen (*preferred stock*) yang memberikan hak dividen tetap dan prioritas klaim aset tetapi umumnya tanpa hak suara (Pinto dkk., 2023). Berdasarkan kinerja dan karakteristik pasar, saham dikategorikan menjadi berbagai tipe seperti saham pertumbuhan (*growth*), pendapatan (*income*), nilai (*value*), dan termasuk *blue chip* (Damodaran, 2022).

Pada periode perdagangan saham, indikator utama yang menunjukkan proses pembentukan harga adalah variabel-variabel Open, High, Low, dan Close. Nilai *Open* merupakan nilai transaksi ketika pasar dibuka. Nilai *Open* menunjukkan ekspektasi investor terhadap informasi yang tersedia sebelum perdagangan dimulai. Pada saat sesi perdagangan berakhir, nilai *Close* dibentuk sebagai nilai transaksi

terakhir, yang menunjukkan keseimbangan pasar setelah pelaku pasar menanggapi semua informasi yang tersedia (Huang et al., 2024). Hubungan antar variabel *Close* dan *Open* dapat diinterpretasikan sebagai apabila *Close* > *Open*, maka permintaan lebih dominan daripada penawaran, sedangkan apabila *Close* < *Open*, maka tekanan jual lebih kuat. Pada penelitian ini dataset yang digunakan adalah data univariate time series, dengan variabe *date* dan *close*.

2.2.2 Saham *Blue chip*

Saham *blue chip* merujuk pada saham perusahaan-perusahaan besar yang memiliki rekam jejak panjang dalam hal stabilitas finansial, profitabilitas konsisten, dan posisi kepemimpinan pasar yang mapan (Pinto dkk., 2023). Istilah *blue chip* secara metaforis berasal dari poker, dimana *chip* berwarna biru (*blue*) memiliki nilai tertinggi.

Saham *blue chip* memiliki karakteristik utama yang meliputi:

1. Ukuran dan Dominasi Pasar, yaitu kapitalisasi pasar yang besar dan merupakan pemain kunci dalam industrinya (Bodie dkk., 2021).
2. Stabilitas Finansial, yaitu memiliki struktur modal yang sehat, arus kas yang kuat, dan riwayat pendapatan yang stabil di berbagai siklus ekonomi (Damodaran, 2022).
3. Dividen yang Konsisten, yaitu memiliki sejarah pembayaran dividen yang teratur dan cenderung meningkat, menjadikannya komponen penting bagi strategi investasi pendapatan (Ang, 2023), dan
4. Reputasi dan tata kelola korporat yang baik, yaitu dikelola oleh perusahaan dengan merek yang kuat dan praktik tata kelola yang transparan (Pompian, 2021).

Pada investasi saham *blue chip* dianggap lebih aman, akan tetapi investasinya tetap saja tidak bebas risiko. Potensi risiko-risiko yang dapat terjadi adalah:

1. Risiko Disrupsi, yaitu perusahaan besar dapat menjadi korban inovasi disruptif dari pesaing yang lebih lincah (Pompian, 2021).

2. Risiko Valuasi, yaitu reputasi sebagai aset yang aman dapat mendorong valuasi yang terlalu mahal (*overvalued*), sehingga membatasi potensi apresiasi masa depan (Damodaran, 2022).
3. Risiko Pertumbuhan Terbatas, yaitu ukuran sangat besar membuat pertumbuhan secara eksponensial menjadi lebih sulit dicapai dibandingkan perusahaan yang lebih kecil (Ang, 2023).
4. Risiko Spesifik Pasar Berkembang, yaitu pada pasar *emerging* seperti Indonesia, klasifikasi dan kinerja *blue chip* dapat dipengaruhi secara signifikan oleh faktor makroekonomi, kebijakan regulasi, dan dinamika kepemilikan institusional yang unik (Sari & Damayanti, 2024).

2.2.3 *Forecasting* (Peramalan)

Forecasting atau peramalan merupakan proses yang digunakan untuk memprediksi kondisi atau nilai suatu variabel di masa depan berdasarkan data historis, pola yang ada, serta teknik analisis tertentu. *Forecasting* bertujuan untuk memberikan gambaran tentang kemungkinan kejadian atau nilai di masa depan dengan tujuan untuk mengurangi ketidakpastian dalam perencanaan dan pengambilan keputusan. *Forecasting* dapat dibagi menjadi dua, yaitu *forecasting* kuantitatif dan *forecasting* kualitatif. Pada *forecasting* kuantitatif, berfokus pada data historis dan statistik untuk membangun model matematis yang memprediksi variabel yang ingin diamati dan bertipe numerik. Sedangkan *forecasting* kualitatif lebih banyak bergantung pada pengalaman, opini ahli, dan informasi yang tidak dapat dijelaskan dengan data numerik. Penelitian ini berfokus pada *forecasting* kuantitatif.

Menurut Hyndman dan Athanasopoulos (2021), *forecasting* dapat digunakan untuk memprediksi berbagai jenis variabel, termasuk permintaan produk, harga komoditas, cuaca, dan tren ekonomi. Dalam implementasinya, *forecasting* memiliki dampak langsung terhadap pengambilan keputusan strategis dalam perencanaan sumber daya, produksi, pemasaran, dan pengelolaan risiko.
Forecasting kuantitatif lebih banyak digunakan ketika data historis yang cukup tersedia. Beberapa teknik umum dalam *forecasting* kuantitatif adalah:

1. Model ARIMA (*Autoregressive Integrated Moving Average*), digunakan untuk meramalkan data yang stasioner dan memiliki pola yang dapat diprediksi.
2. *Exponential Smoothing*, Menggunakan nilai data terbaru untuk memberikan prediksi yang lebih akurat.
3. *Machine Learning*, digunakan untuk mengidentifikasi hubungan antara dua atau lebih variabel dan untuk memprediksi nilai variabel dependen (Hyndman & Athanasopoulos, 2021; Shumway & Stoffer, 2023).

Dalam banyak bidang, *forecasting* menjadi dasar dalam pengambilan keputusan. Dengan *forecasting*, organisasi atau individu akan terbantu dalam merencanakan strategi yang efektif.

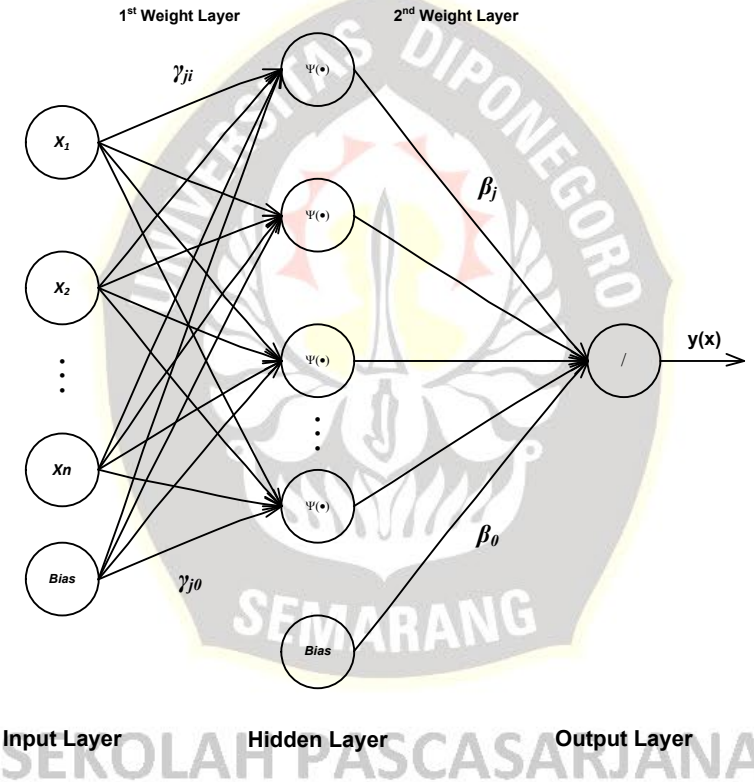
2.2.4 Metode Machine Learning untuk Regresi

Machine Learning (ML) merupakan cabang dari Kecerdasan Buatan (AI) yang memungkinkan komputer untuk mempelajari pola dari data dan membuat prediksi atau keputusan berdasarkan pola tersebut, tanpa pemrograman eksplisit. Dalam *supervised learning*, data pelatihan memiliki label yang digunakan untuk melatih model agar dapat memetakan hubungan antara input (fitur) dan output (target) yang bernilai kontinu untuk permasalahan regresi (Shalev-Shwartz & Ben-David, 2021). Regresi adalah teknik dalam *supervised learning* yang digunakan untuk memprediksi nilai kontinu berdasarkan variabel input yang ada (James dkk., 2021). Sebagai contoh, dalam analisis prediktif ekonomi, model regresi dapat digunakan untuk memprediksi harga saham atau tingkat inflasi berdasarkan data historis yang ada (James dkk., 2021). Metode yang digunakan dalam regresi ML dapat berupa regresi linier, regresi berbasis *Decision tree*, *support vector regression*, k-NN, *Artifial Neural Netwok*, dan lain-lain. Pada penelitian ini, metode *machine learning* yang digunakan adalah model *Neural Network*.

2.2.5 Model *Artificial Neural Network*

Fakta bahwa model ANN memiliki kemampuan untuk menangani masalah non-linier adalah alasan untuk menggunakannya. *Multilayer perceptron* (MLP), juga disebut *feed-forward Neural Network* (FFNN), adalah model ANN yang paling umum. Model MLP terdiri dari berbagai lapisan yang terhubung satu sama lain

dengan parameter atau bobot koneksi, yang menunjukkan kekuatan koneksi. Lapisan tersembunyi berada di antara lapisan input dan output. Neuron di setiap lapisan terhubung dengan bobot fleksibel yang disesuaikan berdasarkan bias. Lapisan input memasukkan data dan melakukan pengolahan data tersembunyi. Lapisan keluaran berfungsi sebagai hasil dari pengolahan data tersebut. Gambar 2.1 memperlihatkan arsitektur MLP yang sangat popular. Aplikasi ANN untuk pemodelan runtun waktu (*time series*) dan pengolahan sinyal biasanya berbasis arsitektur MLP atau FFNN (Suhartono, 2008).



Gambar 2. 1 Arsitektur *Neural Network Backpropagation* (Suhartono, 2008*)*

Nilai y(x) pada model MLP dihitung dengan persamaan sebagai berikut (Suhartono, 2008):

$$y(x) = \beta_0 + \sum_{j=1}^{H} \beta_j \Psi(\gamma_{j0} + \sum_{i=1}^{n} \gamma_{ji} x_i) \,, \tag{2.1}$$

Dimana $\beta_0, \beta_1, \ldots, \beta_H$, dan $\gamma_{10}, \ldots, \gamma_{Hn}$ adalah bobot-bobot dari MLP, n adalah jumlah fitur dan H adalah jumlah hidden layer. Bentuk nonlinear dari fungsi y(x)

terjadi melalui suatu fungsi yang disebut fungsi aktifasi Ψ, yang biasanya fungsi *smooth* seperti fungsi logistik *sigmoid* berikut:

$$\Psi(Z) = \frac{1}{1+e^{-Z}} \tag{2.2}$$

Arsitektur MLP tersusun dari lapisan-lapisan yang disebut layer. Lapisan-lapisan penyusun MLP dibagi mejadi 3, yaitu : *Input Layer*, *Hidden Layer* dan *Output Layer.*

1. **Lapisan Input (*Input Layer*)**

Neuron-neuron dalam lapisan input disebut unit-unit input. Neuron input bertugas menerima inputan dari luar yang menggambarkan fitur-fitur dari permasalahan.

2. **Lapisan Tersembunyi (*Hidden Layer*)**

Neuron-neuron dalam lapisan tersembunyi disebut neuron tersembunyi, yang mana nilai outputnya tidak dapat diamati secara langsung.

3. **Lapisan Output (*Output Layer*)**

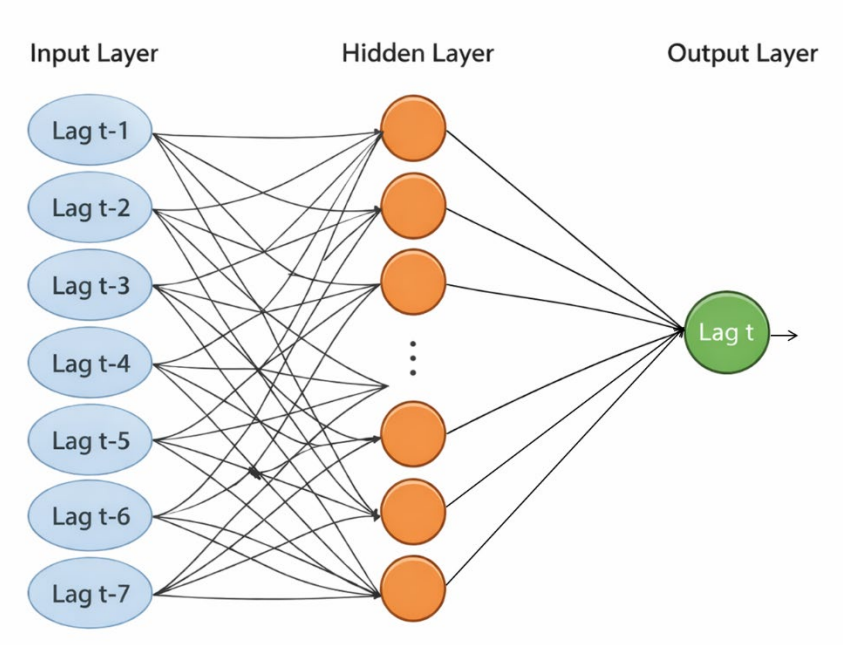
Neuron-neuron dalam lapisan output disebut unit output, yang merupakan solusi dari MLP.

Arsitektur MLP banyak dipakai sebagai baseline pada tugas regresi maupun klasifikasi karena sifatnya yang fleksibel dan mudah diimplementasikan dalam kerangka kerja deep learning (Chollet, 2021; Géron, 2022). Model MLP termasuk dalam kategori *deep learning* apabila arsitekturnya merupakan *deep Neural Network*, yaitu *Neural Network* memiliki *layer* banyak yang memodelkan hubungan nonlinier melalui tumpukan beberapa lapisan tersembunyi (*hidden layers*). *Neural Network* dengan satu hidden layer disebut *shallow Neural Network*, sedangkan *Neural Network* dengan lebih dari satu hidden layer (≥2 hidden layers) dinamakan dengan *deep Neural Network* dan digolongkan sebagai *deep learning* (Sarker, 2021).

2.2.6 Model *Artificial Neural Network* untuk data *univariate time series*

Arsitektur MLP yang digunakan dalam penelitian ini adalah dengan menggunakan 7 neuron pada input layer, yaitu lag y_{t-1} hingga y_{t-7}. Selanjutnya, seluruh input dihubungkan secara penuh (*fully connected*) ke satu *hidden layer* yang

berisi dengan n neuron. Masing-masing *neuron hidden* menghitung penjumlahan berbobot dari seluruh input dan menambahkan bias, yang selanjutnya menerapkan fungsi aktivasi nonlinier. Output layer terdiri dari satu neuron untuk menghasilkan prediksi nilai pada waktu t ($\hat{y}_t$). Arsitektur MLP tersebut dapat diilustrasikan pada gambar 2.2 berikut ini.



Gambar 2. 2 Arsitektur *MLP* dengan 7 input neuron dan 1 ouput neuron

2.2.8 Optimisasi *hyperparameter* dengan metode *Grid Search*

Hyperparameter tuning (*tuning Hyperparameter*) merupakan bagian dari proses *model selection* dengan tujuan untuk meningkatkan kemampuan generalisasi model pada data baru. Pemilihan *hyperparameter* pada metode *machine learning* yang kurang tepat dapat menurunkan kinerja model. *Hyperparameter tuning* diperlukan dalam model *machine learning* agar konfigurasi model lebih stabil (Arnold dkk., 2024).

Metode *Grid Search* merupakan salah satu metode tuning paling umum. Metode *Grid Search* melakukan pencarian yang eksplisit dan menyeluruh pada ruang kandidat hyperparameter yang sudah ditetapkan. Himpunan nilai kandidat

untuk tiap *hyperparameter* ditentukan kemudian membentuk semua kombinasi nilai tersebut, dan mengevaluasi performa setiap kombinasi untuk memilih yang terbaik (Alibrahim & Ludwig, 2021; Bartz-Beielstein & Zaefferer, 2023).

Menurut Alibrahim & Ludwig (2021) dan Bartz-Beielstein & Zaefferer (2023), prosedur ringkas Metode *Grid Search* sebagai berikut:

1. Menentukan *grid hyperparameter*,
2. Untuk setiap kombinasi, latihlah model pada data *training*,
3. Hitung metrik evaluasi pada data validasi, dan
4. Pilih kombinasi yang memberikan skor terbaik sebagai kandidat konfigurasi final.

Metode *Grid Search* banyak dipakai karena memiliki beberapa keunggulan utama, yaitu:

1. Metode *Grid Search,* sederhana dan mudah diimplementasikan. Prosesnya langsung menentukan kandidat nilai, kemudian bentuk kombinasinya dan melakukan evaluasi. Hal ini membuat metode *Grid Search* cocok sebagai metode *tuning hyperparameter* dalam eksperimen penelitian (Bartz-Beielstein & Zaefferer, 2023; Agrawal., 2021).
2. Metode *Grid Search* menjamin pencarian menyeluruh pada grid yang telah ditetapkan. Semua kombinasi pada grid kemudian diuji dan konfigurasi terpilih merupakan yang terbaik di antara kandidat yang didefinisikan. Metode *Grid Search* dapat memberi kepastian dalam ruang pencarian yang telah dibatasi (Alibrahim & Ludwig, 2021).

Ilustrasi metode *Grid Search* adalah sebagai berikut:

Misalkan parameter yang digunakan adalah *Learning rate* dan *Epoch.*

Learning rate = [0.01, 0.1, 1]

Epoch = [50, 100]

Maka metode *Grid Search* akan mencoba seluruh kombinasi parameter tersebut. Seluruh kombinasi tersebut ditunjukan pada tabel berikut:

TABEL 2. 1 KOMBINASI PARAMETER MODEL

No	*Learning Rate*	*Epoch*
1	0.01	50
2	0.01	100
3	0.1	50
4	0.1	100
5	1	50
6	1	100

Setiap kombinasi dilatih dan dievaluasi, kemudian metode *Grid Search* memilih kombinasi yang menghasilkan performa terbaik.

2.2.9 Jenis-Jenis Seleksi Fitur

Seleksi fitur merupakan proses dalam *preprocessing* data yang digunakan untuk model *Machine Learning*. Tujuan seleksi fitur adalah mengurangi dimensi data dengan memilih fitur yang *relevan* dan mengeliminasi fitur yang *tidak informatif*, *redundan*, atau *berisik* (*noise*). Selain itu, tujuan utama seleksi fitur adalah untuk meningkatkan kinerja prediksi dan efisiensi komputasi model, serta mengurangi kemungkinan *overfitting* (Pudjihartono, Aryanti, & Wijaya, 2022).
Dalam seleksi fitur terdapat tiga kategori utama metode seleksi fitur yang digunakan, yaitu:

1. **Metode *Filter***

   Metode ini bekerja secara *independen* terhadap algoritma pembelajaran dengan memilih fitur berdasarkan statistik atau skor relevansi, seperti *correlation* atau *mutual information* (Biernacki, Zbigniew, & Taleb, 2025). Metode Filter memberikan komputasi yang cepat, namun mungkin tidak optimal ketika ada interaksi kompleks antar fitur.

2. **Metode *Wrapper***

   Metode ini melakukan pencarian *subset* fitur dengan mengevaluasi kinerja model pada setiap kombinasi fitur yang berbeda. Keuntungan utama dari metode ini adalah kemampuannya untuk memilih fitur yang memberikan

kinerja terbaik, meskipun biaya komputasinya lebih tinggi (Biernacki dkk., 2025).

3. **Metode *Embedded***

   Seleksi fitur diintegrasikan ke dalam proses pelatihan model itu sendiri, seperti pada regresi *lasso* atau *ridge* yang melibatkan penalti untuk memilih subset fitur yang relevan. Metode *Embedded* menawarkan keuntungan komputasi yang lebih efisien dibandingkan metode *wrapper* (Biernacki dkk., 2025).

Seleksi fitur dalam kasus regresi memiliki tujuan memilih variabel prediktor yang paling berpengaruh terhadap variabel target tanpa memperkenalkan kompleksitas berlebih pada model (Süpürtülü dkk., 2025).

2.2.10 Seleksi Fitur dengan Metode *Forward Selection*

Metode *Forward Selection* merupakan salah satu pendekatan *wrapper feature selection* yang iteratif. Proses utama dari metode ini adalah sebagai berikut:

1. Dimulai dengan tanpa fitur dalam model (*empty set*).
2. Untuk setiap iterasi, satu fitur baru dipilih dari fitur yang tersisa berdasarkan kontribusinya terhadap peningkatan kinerja mode, misalnya pengurangan *error* dalam regresi.
3. Proses dilanjutkan sampai tidak ada peningkatan signifikan pada kinerja model atau sampai mencapai kriteria penghentian (misalnya jumlah fitur maksimum atau kriteria statistik tertentu) (Shafiee, Khayati, & Moghaddam, 2021).

Metode *Forward Selection* sangat populer dalam regresi karena dapat memilih subset fitur yang paling relevan secara bertahap dan mengurangi kemungkinan *overfitting*. Dalam *Machine Learning regresi*, metode *Forward Selection* sering digunakan dengan model regresi yang kuat, seperti *Support Vector Regression* (SVR), untuk memilih fitur yang relevan dalam prediksi hasil atau analisis (Shafiee dkk., 2021).

2.2.11 Evaluasi

Penelitian dapat mendapatkan hasil yang bermacam-macam terutama pada data time series. Hasil tersebut dapat dievaluasi dengan berbagai macam cara. Evaluasi dapat menggunakan persamaan *Root Mean Square Error* (RMSE), *Mean Squared Error* (MSE), dan *Mean Absolute Percentage Error* (MAPE).

Berikut ini merupakan persamaan dari RMSE, MSE dan MAPE untuk evaluasi dalam penelitian ini.

1. RMSE (*Root Mean Square Error*) (Mandal dkk., 2020)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_t - \hat{Y}_t)^2}{n}} \tag{2.6}$$

Dimana Y_t adalah data aktual saham, $\hat{Y}$ adalah hasil prediksi saham dalam data testing sebanyak n.

2. MSE (*Mean Squared Error*) (Mandal dkk., 2020)

$$MSE = \frac{\sum_{i=1}^{n}(Y_t - \hat{Y}_t)^2}{n} \tag{2.7}$$

Dimana Y_t adalah data aktual saham, $\hat{Y}$ adalah hasil prediksi saham dalam data testing sebanyak n.

3. Koefisien determinasi (*coefficient of determination*)

Koefisien determinasi atau R^2 merupakan nilai ukuran statistik yang digunakan untuk menilai model dalam menjelaskan variasi variabel dependen berdasarkan variabel independen. Secara matematis, R^2 dapat dihitung dengan:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{2.8}$$

Dimana SS_{res} adalah jumlah kuadrat residual, dan SS_{tot} adalah jumlah kuadrat total. Nilai R^2 berkisar antara 0 hingga 1, di mana nilai yang lebih tinggi menunjukkan kemampuan model yang lebih baik. Menurut Jian Gao (2024), R^2 merupakan ukuran yang paling umum digunakan dalam regresi untuk

mengevaluasi *goodness of fit*, yaitu sejauh mana model mampu merepresentasikan data yang diamati,