

BAB 2

LANDASAN TEORI

Terdapat beberapa teori-teori dan metode yang digunakan yaitu *Technical Analysis*, *Principal Component Analysis* (PCA), *Support Vector Machine* (SVM), dan *Performance metrics*.

2.1 Technical Analysis

Technical analysis adalah evaluasi sistematis menggunakan harga, volume, dan indikator untuk memprediksi posisi harga di masa depan [11]. *Technical* indikator adalah sebuah alat yang dirancang berdasarkan teori *technical analysis* dalam mencari perubahan signal jual maupun beli [20].

2.1.1 Moving Average Convergence Divergence (MACD)

Moving Average Convergence-Divergence (MACD) merupakan indikator teknikal yang dibangun dengan persamaan yang sama dengan *Exponential moving average* (EMA) [11]. *Exponential moving average* yang digunakan pada MACD adalah EMA 26 hari, EMA 12 hari dan EMA 9 hari[28]. Persamaan *Exponential moving average* adalah sebagai berikut.

$$E[i] = (1 - \alpha)E[i - 1] + \alpha * O[i] \quad (2.1)$$

keterangan :

$E[i]$ = estimasi nilai dari *moving average*

$E[i - 1]$ = nilai terakhir dari *moving average*

α = *smoothing factor* dari 0 hingga 1

$O[i]$ = harga ke-i dari asset

Smoothing factor (α) dapat di tulis dengan persamaan seperti berikut.

$$\alpha = \frac{2}{P+1} \quad (2.2)$$

keterangan :

P = *Period* (hari) dari MACD 12, 26, 9

2.1.2 Exponential Moving Average (EMA)

Exponential Moving Average (EMA) merupakan jenis indikator *moving average* yang menambahkan pembobotan dalam perhitungan terhadap pergenerakan harga tutup. *exponential moving average* (EMA) mengurangi efek lagging yang biasanya terjadi pada simple moving average (SMA) dengan memberikan bobot beban yaitu *smoothing factor* kepada selisih antara harga terkini dengan harga masa lalu, oleh karena itu exponential moving average (EMA) bergerak lebih cepat dibanding *moving average* lainnya. Salah periode waktu yang sering digunakan oleh EMA adalah kurun waktu 40 hari atau EMA-40, hal ini dikarenakan EMA-40 sering digunakan sebagai batas kerugian dalam kegiatan investasi. [29].

$$EMA_{today} = EMA_{yesterday} + \alpha(price_{today} - EMA_{yesterday}) \quad (2.3)$$

Smoothing factor (α) dapat di tulis dengan persamaan seperti berikut.

$$\alpha = \frac{2}{P + 1} \quad (2.4)$$

keterangan :

P = *Period* (hari)

2.1.3 Relative Strength Index (RSI)

RSI Dalam buku Constance Brown merupakan indikator yang digunakan dalam menentukan range dari market naik dan turun [14]. Persamaan RSI dapat dirumuskan sebagai berikut.

$$RSI(n) = \frac{RS_t^n}{1 + RS_t^n}, RS_t^n = \frac{AG_t^n}{AL_t^n} \quad (2.5)$$

AG merupakan *average gain* dan AL merupakan *everage loss* dari periode 14 hari, yang dapat dirumuskan sebagai berikut.

$$AG_{t+1}^n = \frac{AG_t^n * 13 + G_{t+1}^n}{14} \quad (2.6)$$

$$AL_{t+1}^n = \frac{AL_t^n * 13 + L_{t+1}^n}{14} \quad (2.7)$$

keterangan :

AG = Average gain

AI = Average loss
 G = nilai Gain per periode t
 L = nilai Loss per periode t

keterangan :

RSV(n) = nilai dari indikator pada posisi n
 C_t = nilai harga tutup terbaru
 h_n = nilai dari harga tertinggi pada posisi n
 l_n = nilai dari harga terendah pada posisi n
 l_t = nilai dari harga terendah terbaru

2.1.4 On-balance volume (OBV)

$$OBV = OBV_{previous} + Volume, if close > close_{prev} \quad (2.8)$$

$$OBV = OBV_{previous} + Volume, if close = close_{prev} \quad (2.9)$$

$$OBV = OBV_{previous} - Volume, if close < close_{prev} \quad (2.10)$$

keterangan :

OBV = Current OBV level
 $OBV_{previous}$ = previous OBV level
 Vvolume = last trading volume amount

2.2 Principal Component Analysis (PCA)

Principal component analysis digunakan dalam proses *preprocessing* untuk mengecilkan dimensi *input* dan juga mengurangi pengulangan (*redundancy*) dari dataset yang digunakan[30]. PCA di implementasikan dengan menggunakan data *open, high, low, close, Volume, Adj* dan data *technical indicator*. Pada penerapannya PCA menggunakan *eigenvalues* dan *eigenvector* yang didapat dari *matrix* fitur - fitur dataset, sehingga menjadi *principal component factor* seperti persamaan berikut [30].

$$\begin{cases} Y_1 = c_{11}X_1 + \dots + c_{1p}X_p \\ Y_2 = c_{21}X_1 + \dots + c_{2p}X_p \\ \cdot \\ \cdot \\ \cdot \\ Y_p = c_{p1}X_1 + \dots + c_{pp}X_p \end{cases} \quad (2.11)$$

keterangan :

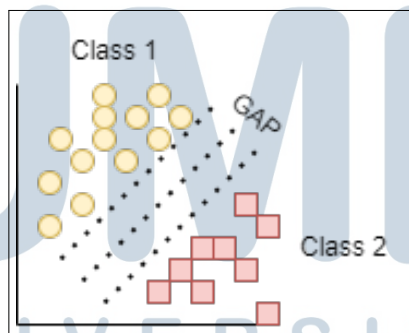
c_{ij} = covariance matrix

X_i = data matrix

Y = principal component factor

2.3 Support Vector Machine (SVM)

Support Vector Machine adalah algoritma machine learning yang berasal dari teori statistika dan dirumuskan oleh Vapnik [22]. SVM membuat batasan keputusan sehingga setiap titik jatuh pada antara batasan yang disebut dengan *hyperplane*.



Gambar 2.1. Hyperplane support Vector machine

Pada Gambar 1 dapat dilihat bahwa SVM bekerja dengan mengklasifikasikan data input kedalam kelas - kelas yang berbeda. *GAP* merupakan margin pada hyperplane yang dapat dirumuskan dengan persamaan sebagai berikut.

$$g(x) = \omega * \phi(x) + b \quad (2.12)$$

ω merupakan vektor dari x , dan b merupakan konstanta bias yang dapat diestimasi dengan *risk function* [31]. $\phi(x)$ adalah fungsi *feature* dengan *input space* x . Persamaan ini bertujuan untuk menentukan maksimum margin dan posisi kelas yang tepat untuk data - data baru.

Persamaan dari *SVM Regularized risk function* adalah sebagai berikut.

$$R(C) = C \frac{1}{N} \sum_{i=1}^n L(d_i, y_i) + \frac{1}{2} \|w^2\| \quad (2.13)$$

keterangan :

C = parameter yang menentukan besar penalti.

N = jumlah data

$L(d_i, y_i)$ = Loss function

Persamaan dari *Loss Function* adalah sebagai berikut.

$$L(d, y) = |d - y| - \epsilon \quad (2.14)$$

keterangan :

d = nilai prediksi

y = nilai yang berasal dari data point

ϵ = konstanta regulasi

Persamaan SVM dengan *Lagrange multipliers* dan *kernel function* adalah sebagai berikut.

$$f(x) = \sum_i^l (\alpha_i - A_i) K(x_i, x_j) + b \quad (2.15)$$

keterangan :

α = konstanta lagrange

A = lagrange multiplier

$K(x_i, x_j)$ = kernel function

b = nilai bias

Persamaan RBF kernel atau yang dikenal sebagai Gaussian kernel adalah sebagai berikut.

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|}{(2\sigma^2)}\right) \quad (2.16)$$

keterangan :

$\|x_i - x_j\|$ = selesih Gaussian euclidean

σ = Gaussian parameter

2.4 Confusion Matrix

Tujuan dari confusion matrix adalah untuk melakukan validasi dan evaluasi dari model yang dirancang. Dalam melakukan evaluasi performa dari SVM dibutuhkan 4 jenis kelas prediksi yaitu True Positive, True Negative, False Positive, False Negative [32]. Hal ini bertujuan untuk menarik kesimpulan kelayakan model untuk diimplementasikan di dunia investasi. Berikut merupakan Table confusion matrix dan persamaan dari *accuracy*, *precision*, *recall* dan F1.

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

Tabel 2.1. Table confusion matrix

$$Accuracy = \frac{TP}{TP + FP + FN + TN} \quad (2.17)$$

$$Precision = \frac{TP}{TP + FP} \quad (2.18)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.19)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (2.20)$$

keterangan :

FP = False Positive (nilai hasil prediksi positif, nilai sebenarnya negatif)

TP = True Positive (nilai hasil prediksi positif, nilai sebenarnya positif)

FN = False Negative (nilai hasil prediksi negatif, nilai sebenarnya positif)

TN = True Negative (nilai hasil prediksi negatif, nilai sebenarnya negatif)



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA