

BAB III METODOLOGI PENELITIAN

3.1 Pengumpulan Dan Seleksi Data

Penelitian ini memakai data harga saham dari PT Catur Sentosa Adiprana yang diperoleh dari laman yahoo.finance.com. Data yang digunakan berupa data harian dari 31 Agustus 2018 hingga 28 Februari 2023. Pada dataset ini terdapat sebanyak 1020 record data dengan 7 atribut yaitu Date, Open, High Low, Close, dan Volume dengan menambahkan beberapa indikator teknikal sebagai berikut :

1. 25MA (*25 simple moving average*)

25 days *simple moving average* dihitung dengan menggunakan rumus rata-rata 25 hari terakhir yaitu:

$$25 MA = \left(\frac{close_t + close_{t-1} + \dots + close_{t-24}}{25} \right)$$

2. 65MA (*65 simple moving average*)

Untuk 65 *simple moving average* menggunakan rumus yang serupa dengan 25 *simple moving average*, yaitu:

$$65 MA = \left(\frac{close_t + close_{t-1} + \dots + close_{t-64}}{65} \right)$$

3. 25EMA (*25 exponential moving average*)

Untuk exponential moving average digunakan rumus berikut untuk mendapatkan nilainya.

$$EMA: \{Close - EMA(\text{previousday})\} \times multiplier + EMA(\text{previousday})$$

4. ALF (*Alexander's filter*)

ALF dapat dihitung dengan rumus:

$$ALF = \left(\frac{close_t}{close_{t-n}} - 1 \right) \times 100$$

5. RS (*Relative Strength*)

Tahapan untuk menghitung Relative Strength adalah sebagai berikut :

- Menghitung Gain/Loss

Jika bernilai positif maka termasuk gain dan jika negatif maka termasuk loss.

$$\text{Gain/loss} = \text{close}_{\text{hari ini}} - \text{close}_{\text{kemarin}}$$

- Menghitung First average Gain/Loss

$$\text{First Average Gain} = \left(\frac{\text{Total Gain 14 hari}}{14} \right)$$

$$\text{First Average Loss} = \left(\frac{\text{Total Loss 14 hari}}{14} \right)$$

- Menghitung Average Gain/Loss

$$\text{Average Gain} = \left(\frac{(\text{Average Gain}_{\text{kemarin}} \times 13) + \text{Gain}_{\text{hari ini}}}{14} \right)$$

$$\text{Average Loss} = \left(\frac{(\text{Average Loss}_{\text{kemarin}} \times 13) + \text{Loss}_{\text{hari ini}}}{14} \right)$$

- Setelah melakukan tahap sebelumnya, barulah kita dapat menghitung relative strength dengan rumus

$$\text{Relative Strength} = \left(\frac{\text{Average Gain}}{\text{Average Loss}} \right)$$

6. RSI (*relative strength index*)

Untuk menghitung RSI diperlukan nilai Relative Strength dari perhitungan sebelumnya.

$$\text{Relative Strength Index} = \left(100 - \frac{100}{1 + \text{Relative Strength}} \right)$$

7. MFI (*money flow index*)

Seperti perhitungan RS sebelumnya, untuk menghitung MFI diperlukan beberapa tahapan perhitungan sebagai berikut:

- Menghitung *Typical Price*

$$\text{Typical Price} = \left(\frac{\text{High} + \text{Low} + \text{Close}}{3} \right)$$

- Menghitung Raw Money Flow

Jika nilai raw money flow lebih besar dibanding raw money flow kemarin, maka Money Flow bernilai Positif dan sebaliknya.

$$\text{Raw Money Flow} = \text{Typical Price} \times \text{Volume}$$

- Menghitung Money Flow Ratio

$$\text{Money Flow Ratio} = \left(\frac{14 \text{ MF positif}}{14 \text{ MF negatif}} \right)$$

- Menghitung Money Flow Index

$$\text{Money Flow Index} = 100 - \left(\frac{100}{1 + \text{Money Flow Ratio}} \right)$$

8. %B Indicator

- Hitung terlebih dahulu nilai 20 days moving average dan 20 days standard deviation.
- Menentukan Upper dan Lower bands.

$$\text{Upper bands} = 20MA + (20stdev \times 2)$$

$$\text{Lower bands} = 20MA - (20stdev \times 2)$$

- Menghitung %B Indikator

$$\%B = \left(\frac{\text{Close} - \text{Lower band}}{\text{Upper band} - \text{Lower band}} \right)$$

9. 10 Days Volatility dan 30 Days Volatility

Untuk menghitung 10 days Volatility dan 30 days volatility cukup menggunakan standar deviasi dengan ketentuan 10 days Volatility = standar deviasi 10 hari, dan 30 days volatility = standar deviasi 30 hari.

10. CHO (Chaikin Oscillator)

Tahapan perhitungan CHO sebagai berikut :

- Menghitung Money Flow Multiplier

$$\text{Money Flow Multiplier} = \left(\frac{(\text{Close} - \text{Low}) - (\text{High} - \text{Low}) - (\text{High} - \text{Close})}{\text{High} - \text{Low}} \right)$$

- Menghitung Money Flow Volume

$$\text{Money Flow Volume} = \text{Money Flow Multiplier} \times \text{Volume}$$

- Menghitung ADL

$$\text{ADL} = \text{ADL (kemarin)} + \text{Money Flow (hari ini)}$$

- Menghitung CHO

$$\text{Chaikin Oscillator} = (3\text{day EMA of ADL} - 10\text{day EMA of ADL})$$

11. MACD (Moving Average Convergence-Divergence)

$$\text{MACD} = 12\text{EMA} - 26\text{EMA}$$

12. %K Indikator dan Williams %R Indicator

Sebelum masuk ke rumus perhitungan, tentukan terlebih nilai tertinggi harga *High* dan nilai terendah harga *Low* dari data yang diobservasi, setelah itu hitung %K indikator dan %R Indicator menggunakan rumus :

$$\%K = \left(\frac{(\text{Close} - \text{Low terendah})}{(\text{High tertinggi} - \text{Low terendah})} \right) * 100$$

$$\%R = \left(\frac{(\text{High tertinggi} - \text{Close})}{(\text{High tertinggi} - \text{Low terendah})} \right) * 100$$

13. Accumulation and distribution (AD) oscillator

AD Oscillator sebelumnya sudah dihitung pada bagian *Chaikin Oscillator*.

Data historis data saham yang telah diperoleh sebelumnya akan dihitung dengan lima belas indikator teknik diatas. Sebelum dilakukan perhitungan menggunakan indikator teknikal, data akan melalui proses cleansing data terlebih dahulu. Pada proses ini data harian harga saham akan dibuang jika bernilai *null*. Selain itu, data

dengan nilai volume = 0 juga akan dibersihkan mengingat adanya beberapa indikator teknikal yang menggunakan volume pembelian dalam proses perhitungannya.

3.2 Pembagian Kelas Data

Sebelum memasuki proses SVM, data harian harga saham akan dibagi kedalam tiga kelas, yaitu Downtrend, Notrend dan Uptrend. Perhitungan pembagian kelas dilakukan pada msExcel dengan menerapkan aturan sebagai berikut :

- a. **Uptrend.** Data diklasifikasikan sebagai uptrend jika memenuhi syarat berikut :
 - Harga Close aktual $>$ nilai 25-day moving average.
 - Nilai 25-day moving average $>$ nilai 65-day moving average.
 - Nilai 25-day moving average harus naik setidaknya untuk 5 hari terakhir.
 - Nilai 65-day moving average harus naik setidaknya untuk 1 hari terakhir.
- b. **Downtrend.** Data diklasifikasikan sebagai downtrend jika memenuhi syarat berikut :
 - Nilai close aktual $<$ nilai 25-day moving average.
 - Nilai 25-day moving average $<$ dari 65-day moving average.
 - Nilai 25-day moving average harus turun setidaknya untuk 5 hari terakhir.
 - Nilai 65-day moving average harus turun setidaknya untuk 1 hari terakhir.
- c. **No trend.** Data akan diklasifikasikan sebagai *no trend* jika data tersebut tidak memenuhi syarat-syarat yang disebutkan sebelumnya.

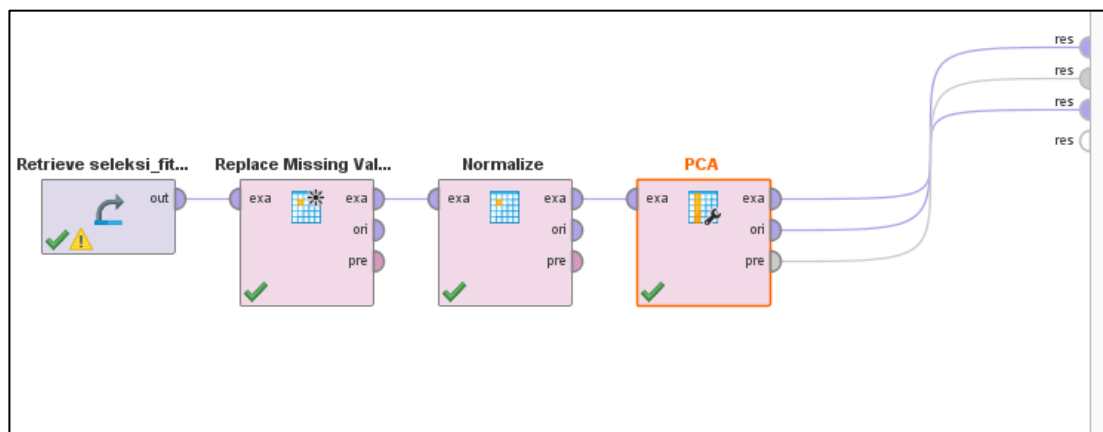
Berikut ini tampilan data harga saham yang telah melalui proses seleksi data, perhitungan fitur serta klasifikasi kelas data.

Date	CLASS	Open	High	Low	Close	Volume	10_day	30_day	25MA	65MA	%B
06/12/2011	Downtrend	208.486	208.486	203.522	208.486	1060142	4.19	9.01	222.98	227.93	-0.87
07/12/2011	Downtrend	213.45	213.45	203.522	213.45	1617157	4.80	9.37	222.39	227.32	-0.73
08/12/2011	Downtrend	213.45	213.45	203.522	213.45	632056	3.92	9.48	222.39	226.63	-0.71
09/12/2011	Downtrend	203.522	208.486	203.522	203.522	1628741	3.66	9.59	222.19	225.79	-0.96
12/12/2011	Downtrend	203.522	208.486	198.558	203.522	5038823	4.56	10.15	221.59	224.95	-0.94
13/12/2011	Downtrend	198.558	203.522	198.558	198.558	2576573	5.34	10.65	221.00	223.96	-1.08
14/12/2011	Downtrend	203.522	203.522	198.558	203.522	196919	6.10	11.36	220.20	223.04	-0.87
15/12/2011	Downtrend	198.558	203.522	198.558	198.558	757963	6.19	11.69	219.01	222.20	-0.99
16/12/2011	Downtrend	208.486	213.45	203.522	208.486	2322743	6.64	12.29	217.02	221.29	-0.63
20/12/2011	Notrend	218.414	238.27	218.414	218.414	4049189	5.36	12.44	215.63	220.37	-0.29
21/12/2011	Notrend	223.378	228.342	218.414	223.378	1091870	6.64	12.44	214.84	219.68	-0.10
22/12/2011	Notrend	223.378	223.378	213.45	223.378	1377933	8.44	12.47	214.44	219.15	-0.12
23/12/2011	Notrend	233.306	238.27	223.378	233.306	3698662	9.59	12.19	214.24	218.69	0.20
27/12/2011	Notrend	233.306	233.306	223.378	233.306	1190582	12.20	11.19	214.64	218.46	0.08
28/12/2011	Notrend	223.378	238.27	223.378	223.378	1848827	13.60	10.54	214.84	218.77	-0.25
29/12/2011	Notrend	223.378	228.342	223.378	223.378	1364335	13.28	9.85	214.84	218.69	-0.27
30/12/2011	Notrend	228.342	228.342	223.378	228.342	152096	11.80	9.44	215.04	218.77	-0.16
02/01/2012	Notrend	228.342	243.234	223.378	228.342	5868806	10.77	9.44	215.24	218.69	-0.19
03/01/2012	Notrend	228.342	233.306	223.378	228.342	2088555	7.33	9.61	215.63	218.77	-0.22
04/01/2012	Notrend	228.342	233.306	223.378	228.342	2165107	4.80	9.61	216.03	218.84	-0.24
05/01/2012	Notrend	223.378	228.342	218.414	223.378	1092374	3.92	9.77	216.63	218.99	-0.37
06/01/2012	Notrend	223.378	228.342	218.414	223.378	326856	3.92	9.84	216.83	219.22	-0.38
09/01/2012	Notrend	223.378	228.342	218.414	223.378	564569	3.92	9.84	217.22	219.64	-0.39
10/01/2012	Notrend	218.414	223.378	213.45	218.414	1629748	3.47	9.90	217.62	219.94	-0.51
11/01/2012	Notrend	218.414	223.378	218.414	218.414	596802	3.35	9.90	217.62	220.09	-0.53
12/01/2012	Uptrend	223.378	228.342	218.414	223.378	1785873	3.92	9.88	218.02	220.25	-0.43

Gambar 3.1 Sampel Data Harga Saham

3.3 Principal Component Analysis

Lima belas indikator sebelumnya akan diseleksi menggunakan Principal Component Analysis, hal ini dimaksudkan untuk memperbesar nilai akurasi prediksi melalui tambahan fitur yang paling berpengaruh terhadap pergerakan harga saham. Proses PCA dilakukan dengan Rapidminer .



Gambar3.2 Proses PCA pada RapidMiner

10_days	30_days	25MA	65MA	%B	%K	%R	RS	RSI	ALF	MACD	25EMA	AD	10_days	30_days
16.3724	45.08485	610.8	598.4615	-0.50922	80	20	0.875	46.66667	1.694915	-5.39105	607.2879	0	96.75499	88.36457
16.3724	43.02365	606.8	599.2308	-1.24419	-80	180	0.636364	38.88889	-5.08475	-8.34182	603.6504	-161800	-49685.4	-55027.2
21.21975	41.84811	601.4	599.6154	-0.56782	77.77778	22.22222	0.954545	48.83721	0.847458	-7.7666	602.985	46080	14274.08	15784.78
21.23938	38.59456	598	600.1538	-1.16473	18.18182	81.81818	0.724138	42	-5.88235	-10.0194	599.6784	0	94.74293	86.48286
16.90661	36.60326	595.8	600.6154	-1.0428	20	80	0.724138	42	-5.08475	-11.6703	596.6262	0	94.6715	86.53141
18.77498	34.32812	595.4	601.0769	-0.89867	30	70	0.758621	43.13725	-4.23729	-12.4319	594.1935	30200	9386.094	10373.45
18.41648	29.99808	593.8	601.4615	-0.4231	100	0	1	50	1.694915	-10.0948	594.6401	2236600	688279.3	762005.1
19.30026	24.09333	594.6	602.4615	-0.4157	100	0	0.689655	40.81633	2.564103	-8.14878	595.0524	-800	-151.789	-186.341
19.69207	18.35239	595.4	603.3846	-0.4157	66.66667	33.33333	1	50	3.448276	-6.53123	595.433	0	98.24647	90.52604
19.69207	17.97508	595.4	604.0769	-0.47814	58.33333	41.66667	0.904762	47.5	2.586207	-5.58835	595.3997	0	98.23112	90.53372
19.30026	17.39418	595.2	604.7692	-0.91332	0	100	0.826087	45.2381	-4.2735	-7.57797	592.6766	0	97.72227	90.11512
19.64264	18.39853	593.2	604.8462	-0.5118	50	50	0.956522	48.88889	0	-6.65726	592.4707	-109500	-33595	-37212.6
18.59659	18.23852	592.4	605.3846	-0.50593	50	50	0.913043	47.72727	0.854701	-5.86005	592.2807	-2300	-610.022	-693.541
18.28782	18.07074	592	605.9231	-0.5	75	25	0.913043	47.72727	5.357143	-5.16866	592.1052	0	97.76431	90.03047
16.66667	18.04847	591.6	606.4615	-0.38259	100	0	0.956522	48.88889	6.25	-4.16922	592.3279	0	97.71083	89.94811
14.34689	18.01101	591.4	606.9231	-0.45409	85.71429	14.28571	1.375	57.89474	0.854701	-3.73753	592.1488	171300	52806.76	58446.42
11.73788	17.74743	591	607.3846	-0.44067	88.88889	11.11111	0.9375	48.3871	0	-3.35672	591.9836	0	97.75531	90.12277
11.30388	17.45355	590.6	607.3846	-0.44842	88.88889	11.11111	1.666667	62.5	2.608696	-3.02011	591.831	0	100.6804	93.28439
10.74968	17.39418	590.2	607.3846	-0.52396	77.77778	22.22222	1.5	60	1.73913	-3.12082	591.3055	-112400	-34484.1	-38197
10.0692	17.3669	587.8	607.1538	-0.59379	66.66667	33.33333	1.272727	56	0	-3.56303	590.4359	540600	166438.4	184253
9.944289	17.40657	587	607.3077	-0.57918	75	25	0.636364	38.88889	3.571429	-3.86888	589.6331	0	97.35522	89.83519
4.830459	17.41977	586	607.6923	-0.47858	75	25	0.727273	42.10526	0.862069	-3.66556	589.2767	0	98.13782	90.78641

Gambar 3.3 Sampel Indikator Teknikal sebagai Input pada proses PCA

Dari proses PCA diperoleh nilai eigenvalues sebagai berikut :

Tabel 3.1 Nilai Eigenvalue

Component	Standard Deviation	Proportion of Variation	Cumulative Variation
PC 1	0.408	0.583	0.583
PC 2	0.254	0.227	0.810
PC 3	0.145	0.074	0.884
PC 4	0.122	0.053	0.936
PC 5	0.073	0.019	0.955
PC 6	0.060	0.013	0.968
PC 7	0.057	0.011	0.979
PC 8	0.050	0.009	0.988
PC 9	0.037	0.005	0.993
PC 10	0.035	0.004	0.997
PC 11	0.020	0.001	0.999
PC 12	0.014	0.001	0.999
PC 13	0.012	0.000	1.000
PC 14	0.002	0.000	1.000
PC 15	0.001	0.000	1.000

Proses PCA akan mengelompokkan lima belas data fitur kedalam lima belas Principal Component(PC). Umumnya reduksi fitur dapat dipilih melalui komponen yang mencapai variansi kumulatif 0.95. Namun dari hasil eigenvalue terlihat banyak komponen yang proporsinya banyak tidak lebih dari 0.1, kecuali PC1 (0.583) dan PC2(0.227) dengan variansi kumulatif 0.810. Dengan demikian

hanya komponen pada PC1 dan PC2 yang akan digunakan selanjutnya dengan anggapan PC lainnya tidak begitu berpengaruh terhadap data.

Tabel 3.2 Nilai Eigenvector

Nama Faktor/Variabel	PC1	PC2
10 Days Volatility	-0.186	0.179
30 Days Volatility	-0.183	0.142
25 Moving Average	-0.551	0.042
65 Moving Average	-0.549	0.003
%B Indicator	0.008	0.103
%K Indicator	0.001	0.027
William %R Indicator	-0.001	-0.027
Relative Strength	0.001	0.157
Relative Strength Index	0.020	0.513
Alexander's Filter	-0.002	0.008
Moving Average Convergence-Divergence	-0.016	0.237
25 Exponential Moving Average	-0.553	0.060
Accumulation and Distribution Oscillator	0.004	0.007
Money Flow Index	0.144	0.766
Chaikin Oscillator	0.001	-0.001

3.4 Prediksi Menggunakan Support Vector Machine

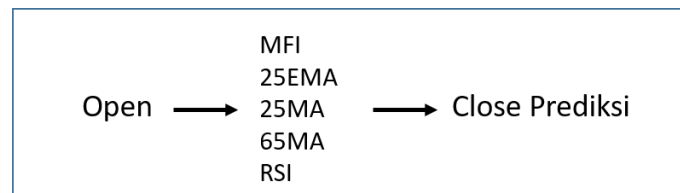
Pada penelitian ini, algoritma Support Vector Machine akan digunakan untuk memprediksi nilai *close* harga saham serta kelas data saham akan dijadikan sebagai penilai tingkat akurasi prediksi. Sebagai perbandingan, proses prediksi akan dibedakan menjadi prediksi PCA-SVM untuk proses SVM yang melibatkan 5 fitur indikator teknikal terpilih serta prediksi SVM normal tanpa tambahan fitur.

Pada prediksi nilai, penulis akan menggunakan nilai open sebagai input utama. Pemilihan nilai open didasarkan pada asumsi penulis terhadap kondisi pasar saham yang mana nilai open merupakan pembuka dari harga suatu saham hari itu, maka secara otomatis nilai open merupakan nilai paling awal yang akan dimunculkan pada website informasi harga saham sehingga hasil prediksi dapat digunakan sebelum nilai close atau penutupan harga saham hari itu diumumkan.

3.4.1 Prediksi PCA-SVM 5 fitur

1. Prediksi Close

Prediksi nilai close akan menggunakan konsep regresi SVM, yang mana memang umum digunakan untuk memprediksi harga saham maupun komoditas lain. Adapun alur prosesnya sebagaiberikut :

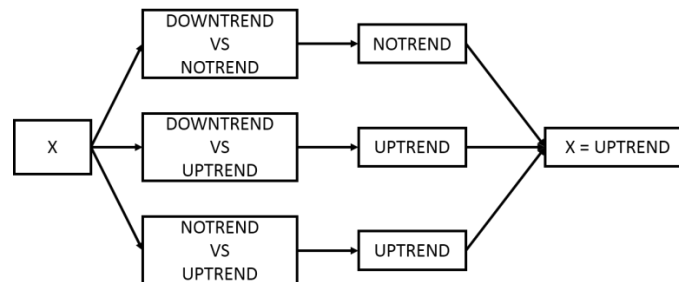


Gambar 3.4 Alur prediksi nilai Close PCA-SVM

Dari proses tersebut, diperoleh beberapa persamaan fungsi yang nantinya akan digunakan untuk

a. Prediksi Class

SVM seperti teknik klasifikasi lainnya pada mulanya hanya digunakan dengan tujuan memisahkan data kedalam dua kelas berbeda. Pada penelitian ini, data akan dikelompokkan ke dalam tiga kelas. Penulis menggunakan metode One vs One dengan alur prediksi seperti pada gambar berikut



Gambar 3.5 Alur Prediksi SVM Multikelas

3.5 Analisa Performa Metode PCA-SVM

Setelah melakukan tahap prediksi langkah selanjutnya adalah evaluasi untuk mengukur ketepatan dan keakuratan metode yang digunakan maka ditentukan dengan menghitung nilai Mean Square Error (MSE), Root Mean Square Error (RMSE), dan Mean Absolute Error (MAE) yang dijelaskan berikut ini:

1. MSE adalah cara yang digunakan pada proses evaluasi suatu prediksi dengan cara mengkuadratkan nilai selisih antara nilai eror actual dan nilai eror prediksi, semakin besar nilai erornya tentu makin besar pula MSE yang dihasilkan. Formula MSE didefinisikan sebagai berikut :

$$MSE = \sum \frac{(Y' - Y)^2}{n}$$

2. RMSE ialah akar dari nilai rata-rata (mean) eror kuadrat, yang diambil dari nilai selisih antara nilai sebenarnya dan nilai prediksi. Formula RMSE didefinisikan sebagai berikut :

$$RMSE = \sqrt{\sum \frac{(Y' - Y)^2}{n}}$$

3. MAE menunjukkan nilai rata-rata kesalahan (error) dari nilai sebenarnya dengan nilai peramalan. Formula RMSE didefinisikan sebagai berikut :

$$MAE = \sum \frac{|Y' - Y|}{n}$$

Diketahui :

Y' : nilai peramalan

Y : nilai sebenarnya

n : jumlah data

3.5 Implementasi Hasil Prediksi

Setelah ditemukan metode prediksi PCA-SVM terbaik dari evaluasi yang dilakukan, selanjutnya hasil prediksi tersebut akan diimplementasikan ke dalam sebuah website sederhana berbentuk dashboard.

