

Prediksi Trend Pergerakan Harga Saham dengan *Hidden Markov Model (HMM)* dan *Support Vector Machine (SVM)*

Firdaniza¹, Jondri²

¹Prodi Matematika, Fakultas MIPA, Universitas Padjadjaran
Jl. Raya Bandung Sumedang KM 21 Jatinangor Sumedang 45363
Email: firdaniza.unpad@gmail.com

²Prodi Teknik Informatika Universitas Telkom
Jl. Telekomunikasi, Terusan Buah Batu Bandung
Email: jdn@ittelkom..ac.id

ABSTRAK

Prediksi trend pergerakan harga saham sangatlah dibutuhkan untuk meningkatkan potensi keuntungan sekaligus mengurangi kemungkinan rugi. Berbagai metode telah digunakan untuk memprediksi trend pergerakan harga saham. Pada paper ini, dibahas metode *Hidden Markov Model (HMM)* dan *Support Vector Machine (SVM)* sebagai alat untuk memprediksi trend naik turunnya harga *close* Indeks LQ45. Akurasi prediksi dengan HMM sebesar 50,98%, sementara dengan SVM sebesar 55,56%.

Kata kunci: *Hidden Markov Model, Support Vector Machine, Harga Saham.*

ABSTRACT

Predicted trend of stock price movements are needed to increase the profit potential while reducing the possibility of loss. Various methods have been used to predict the trend of stock price movements. In this paper, we use Hidden Markov Models (HMM) and Support Vector Machine (SVM) as a tool to predict trend of stock price fluctuations of the close price of LQ45 index. With HMM prediction accuracy of 50.98%, while the SVM was 55.56%

Keywords: *Hidden Markov Model, Support Vector Machine, stock price.*

1. Pendahuluan

Saham adalah satuan nilai dalam berbagai instrument finansial yang mengacu pada bagian kepemilikan sebuah perusahaan. Dengan menerbitkan saham, memungkinkan perusahaan-perusahaan yang membutuhkan dana jangka panjang untuk 'menjual' dengan imbalan uang tunai untuk modal bisnisnya. Harga saham dapat dilihat sewaktu-waktu di bursa efek kapanpun orang menginginkan. Investor akan memperhatikan dinamika harga saham setiap saat untuk memperkirakan kapan dia harus menjual atau membeli saham agar dia memperoleh keuntungan. Indeks LQ45 merupakan indeks harga saham dari 45 saham pilihan dengan tingkat likuiditas dan kapitalisasi pasar tertinggi. Indeks LQ45 bertujuan sebagai pelengkap IHSG dan khususnya menyediakan sarana yang objektif dan terpercaya bagi analis keuangan, manajer investasi, investor, dan pemerhati pasar modal lainnya dalam memonitor pergerakan harga dari saham-saham yang aktif di perdagangan. Untuk dapat memprediksi trend harga saham, investor harus mempunyai teknik analisis yang kuat sebagai dasar pengambilan keputusan. Berbagai metoda sudah banyak digunakan.

Untuk memprediksi harga saham atau trend harga saham dapat digunakan analisis fundamental dan analisis teknikal. Analisis fundamental menggunakan data tentang laporan keuangan perusahaan, manajemen perusahaan, kondisi makro ekonomi, serta kondisi mikro ekonomi terkait dengan bidang usaha perusahaan yang akan diprediksi harga sahamnya. Analisis teknikal menggunakan data historis harga saham suatu perusahaan.

Baresa [1] menganalisa pengaruh makro dan mikro ekonomi pada bisnis perusahaan untuk memprediksi harga saham perusahaan tersebut di masa mendatang. Venkates [6] menggunakan data tentang kebijakan keuangan, keuntungan, dan efisiensi perusahaan farmasi, keuangan, dan teknologi informasi di bursa saham India untuk memprediksi *return* dari saham perusahaan tersebut, dimana *return* ini akan menjadi faktor penyebab naik atau turunnya harga saham suatu perusahaan.

Analisis teknikal yang biasa digunakan oleh para pialang saham adalah *moving average*, karena metodanya paling sederhana. Harga saham besok diramal dari harga rata-rata harga saham hari-hari sebelumnya. Hasan-Nath [3], menggunakan *Hidden Markov Model* (HMM) untuk memprediksi harga saham. Data yang digunakan adalah harga *close*.

Pada paper ini dikaji penerapan HMM untuk memprediksi trend harga saham dengan menggunakan harga *close*. Disamping itu, supaya dapat menggunakan data histori lebih banyak yang diformulasikan dalam berbagai macam teknikal indikator, akan dikaji metoda *Support Vector Machine* (SVM) untuk memprediksi trend harga saham dengan menggunakan beberapa teknikal indikator sebagai atribut.

2. Metode Penelitian

Untuk kasus data harga saham, state tidak dapat diamati secara langsung, tetapi dapat diobservasi melalui himpunan pengamatan lain, dalam hal ini data harga saham berbentuk *open, high, low, close*. Dalam metoda ini, satu set data masa lalu dianalisis dan dilakukan training *HMM* yang selanjutnya digunakan untuk memprediksi trend harga saham hari yang akan datang. Pada paper ini akan digunakan *Hidden Markov Model* (HMM) dan *Support Vector Machine* (SVM) dalam memprediksi trend pergerakan harga indeks LQ45. Nilai *close* dari indeks LQ45 diklasifikasi ke dalam “naik” dan “turun”. Sebagian data dijadikan data *training* dan sebagian lainnya dijadikan data *testing*. Dengan HMM dan SVM akan dilihat akurasi dalam memprediksi trend harga saham.

3. Hasil dan Pembahasan

3.1 Hidden Markov Model (HMM)

Hidden Markov Model (HMM) adalah rantai *Markov* dengan *statenya* tidak dapat diamati secara langsung, tetapi hanya dapat diobservasi melalui himpunan pengamatan lain, Elemen-elemen *HMM* antara lain: Rabiner [5]

1. X_t yaitu *state* pada waktu t .
2. $X = \{X_1, X_2, \dots, X_N\}$ yaitu ruang *state* dengan N adalah jumlah *state*.
3. O_t yaitu observasi pada waktu t .
4. $V = \{v_1, v_2, \dots, v_M\}$ yaitu ruang observasi dengan M adalah jumlah observasi.
5. π_i yaitu distribusi *state* awal.
6. $\mathbf{A} = [a_{ij}]$, yaitu matriks peluang transisi dari *state* i ke j .
7. $\mathbf{B} = [b_{jm}]$, yaitu matriks peluang observasi v_m pada *state* j , dimana:
8. $b_{jm} = b_j(O_t) = P(O_t = v_m | X_t = s_j)$, $1 \leq j \leq N$ dan $1 \leq m \leq M$
9. *HMM* dinotasikan dengan $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$

Prediksi trend harga saham dengan HMM dilakukan dengan menyelesaikan 3 masalah utama HMM, yakni :

- a. *Evaluation problem*
- b. *Decoding problem*
- c. *Learning problem*.

Pada *evaluation problem*, dihitung $P(O|\lambda)$ dengan prosedur *forward*, artinya dicari peluang barisan observasi jika diberikan parameter $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$. Dalam hal ini digunakan variabel *forward*,

$$\alpha_t(i) = P(O_1, O_2, \dots, O_t, X_t = S_i | \lambda)$$

Selanjutnya, pada *decoding problem*, ditentukan barisan *state* optimal $X = \{X_1, X_2, \dots, X_N\}$ jika diberikan parameter $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ dan barisan observasi O . *Decoding problem* diselesaikan dengan memanfaatkan variabel *forward* dan *backward*, yakni $\gamma_t(i) = P(X_t = s_i | O, \lambda)$.

Kemudian *learning problem*, menentukan parameter HMM $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ agar diperoleh $P(O|\lambda)$ maksimum.

Metode yang digunakan dalam learning problem ini adalah algoritma Baum-Welch. Pada algoritma ini digunakan variabel *forward*, *backward*, $\gamma_t(i)$, dan $\xi_t(i, j)$.

$$\xi_t(i, j) = P(X_t = s_i, X_{t+1} = s_j | O, \lambda)$$

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j)$$

$P(O | \lambda)$ akan maksimum jika $\sum_{t=1}^{T-1} \gamma_t(i) \neq 0$.

Langkah-langkah prediksi trend harga *close* indeks LQ45 dengan menggunakan HMM adalah sebagai berikut:

1. Klasifikasikan nilai *close* harga saham LQ45 periode 4 Maret 2013 sampai 27 Januari 2014 menjadi "naik" dan "turun" dengan mencari selisih harga *close* hari ini dengan harga *close* hari kemarin. Diperoleh data sebanyak 210 data, dibagi menjadi data *training* sebanyak 160 buah dan 50 sebagai data *testing*.
2. Gunakan data *training* uuntuk memperoleh matriks peluang transisi A dan matriks peluang emisi B dengan Algoritma Baum-Welch.
Diperoleh :

$$A = \begin{bmatrix} 0,6007 & 0,3993 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0,5567 & 0,4433 \\ 0,4567 & 0,5433 \end{bmatrix}$$
3. Dengan HMM *decoding*, hitung peluang observasi $O = \{O_{d-9}, O_{d-8}, \dots, O_d, O_{d+1}\}$ dengan O_{d+1} adalah kemungkinan nilai observasi (naik/turun) . Artinya, dari 10 nilai *close* sebelumnya, dihitung peluang hari ke 11 akan "naik" atau "turun".
4. Harga saham hari ke-11 diprediksi "naik" jika peluang harga "naik" lebih dari peluang harga "turun", sebaliknya, harga saham hari ke-11 diprediksi "turun" jika peluang harga "turun" lebih dari peluang harga "naik"
5. Bandingkan hasil prediksi dengan data sebenarnya, diperoleh akurasi 50,98%.

3.2 Support Vector Machine (SVM)

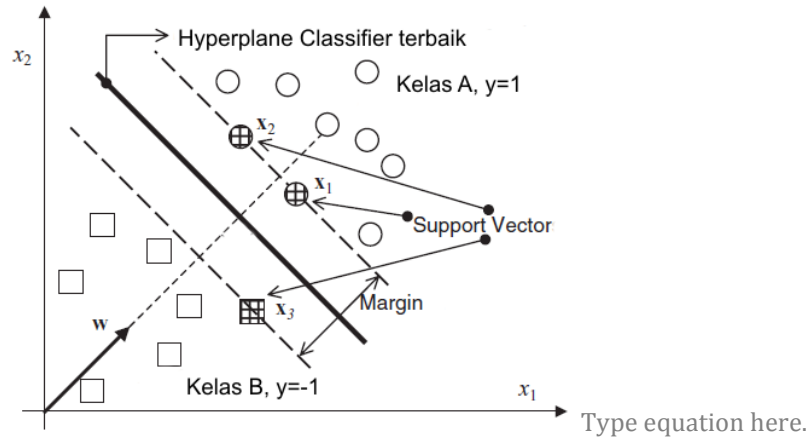
Support Vector Machine (SVM) adalah metode yang digunakan untuk mengklasifikasikan data, diterapkan pada banyak permasalahan nyata seperti pada *data mining*, *machine learning* dan pengenalan pola, Cristianini [2]. Prinsip kerja dari metoda SVM adalah menentukan model *classifier* (biasanya berupa *hyperplane*) yang memisahkan data dari dua kelas yang berbeda. Koefisien-koefisien dari *hyperplane classifier* dipilih sehingga jarak *hyperplane* ke titik-titik terdekat dari kedua kelas paling jauh.

Diberikan himpunan data $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N) \in R^n \times R$ yang dapat dipisah secara linear ke dalam kelas A (dinotasikan dengan indeks 1) dan kelas B (dinotasikan dengan indeks -1). Dalam hal ini x_i menyatakan atribut.

Hyperplane terbaik yang memisahkan kedua data tersebut dapat diperoleh berupa *hyperlane*

$$w^T x + b = 0 \tag{1}$$

Sehingga jarak *hyperplane* ke titik terdekat dari kedua kelas bernilai maksimal (Ilustrasi dalam dua dimensi dapat dilihat pada Gambar 1).



Gambar 1. Ilustrasi *hyperplane classifier* terbaik [2]

Agar *hyperplane* (1) memisahkan data ke dalam kedua kelas, maka dapat dipilih vektor \mathbf{w} dan \mathbf{b} sehingga

$$\mathbf{w}^T \mathbf{x}_i + \mathbf{b} > 0, \quad \text{jika } \mathbf{x}_i \in A \tag{2}$$

dan

$$\mathbf{w}^T \mathbf{x}_i + \mathbf{b} < 0, \quad \text{jika } \mathbf{x}_i \in B \tag{3}$$

Untuk suatu titik \mathbf{x}_0 yang diberikan, jarak *hyperplane* ke titik tersebut diberikan oleh

$$J = \frac{|\mathbf{w}^T \mathbf{x}_0 + \mathbf{b}|}{\|\mathbf{w}\|_2} \tag{4}$$

Jarak *hyperplane* ke titik terdekat dari kedua kelas disebut *margin*. Pencarian *hyperplane classifier* dilakukan dengan langkah-langkah berikut.

Andaikan \mathbf{x}^+ dan \mathbf{x}^- masing-masing adalah titik pada himpunan A dan B yang terdekat ke *hyperplane*. Maka tanpa mengurangi keumuman, dapat ditetapkan

$$\mathbf{w}^T \mathbf{x}^+ + \mathbf{b} = 1 \tag{5}$$

dan

$$\mathbf{w}^T \mathbf{x}^- + \mathbf{b} = -1 \tag{6}$$

Sehingga

$$\mathbf{w}^T \mathbf{x}_i + \mathbf{b} \geq 1, \quad \text{untuk semua } \mathbf{x}_i \in A \tag{7}$$

dan

$$\mathbf{w}^T \mathbf{x}_i + \mathbf{b} \leq -1, \quad \text{untuk semua } \mathbf{x}_i \in B \tag{8}$$

Jarak antara *hyperplane* dengan titik \mathbf{x}^+ dan \mathbf{x}^- dapat dituliskan menjadi

$$J = \frac{|\mathbf{w}^T \mathbf{x}^+ + \mathbf{b}|}{\|\mathbf{w}\|_2} = \frac{|\mathbf{w}^T \mathbf{x}^- + \mathbf{b}|}{\|\mathbf{w}\|_2} = \frac{1}{\|\mathbf{w}\|_2} = \frac{1}{\langle \mathbf{w}, \mathbf{w} \rangle^{1/2}} \tag{9}$$

Definisikan

$$y_i = h(\mathbf{x}_i) = \begin{cases} 1, & \text{untuk } \mathbf{x}_i \in A \\ -1, & \text{untuk } \mathbf{x}_i \in B \end{cases} \tag{10}$$

Maka $(\mathbf{w}^T \mathbf{x}_i + \mathbf{b})y_i \geq 1, i = 1, 2, \dots, N$

Koefisien *hyperplane classifier* diperoleh dari solusi yang masalah optimasi berkendala

$$\begin{aligned} & \text{Maks}_{w,b} \frac{1}{\langle \mathbf{w}, \mathbf{w} \rangle^{1/2}} \\ & \text{s.t.} \end{aligned} \tag{11}$$

$$(\mathbf{w}^T \mathbf{x}_i + b) y_i \geq 1, \text{ untuk } i = 1, 2, \dots, N$$

Persamaan ini ekuivalen dengan solusi masalah optimasi

$$\begin{aligned} & \text{Min}_{w,b} \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ & \text{s.t.} \end{aligned} \tag{12}$$

$$(\mathbf{w}^T \mathbf{x}_i + b) y_i \geq 1, \text{ untuk } i = 1, 2, \dots, N$$

Dengan pendekatan metode pengali Lagrange, solusi masalah optimasi berkendala (12) dicari dari solusi masalah *dual*

$$\begin{aligned} & \text{Maks}_{\alpha} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T \mathbf{x}_j \\ & \text{s.t.} \end{aligned} \tag{13}$$

$$\sum_{i=1}^N \alpha_i y_i = 0, \quad \alpha_i \geq 0, i = 1, 2, \dots, N$$

Cara diatas dapat diperluas untuk klasifikasi multikelas dengan cara membandingkan 1 kelas dengan kelas sisanya yang lain [4].

Langkah-langkah prediksi trend harga *close* saham dengan SVM adalah sebagai berikut:

1. Klasifikasikan nilai *close* harga saham LQ45 periode 4 Maret 2013 sampai 27 Januari 2014 menjadi “naik” dan “turun” dengan mencari selisih harga *close* hari ini dengan harga *close* hari kemarin. Diperoleh data sebanyak 210 data, dibagi menjadi data *training* sebanyak 120 buah dan 90 sebagai data *testing*.
2. Atribut yang digunakan adalah teknikal indikator berikut:
 Stokastik %K: $\frac{C_t - LL_n}{HH_n - LL_n}$, stokastik %D: $\frac{\sum_{i=0}^{n-1} \%C_{t-i}}{n}$, momentum: $C_t - C_{t-4}$, Rate of Change: $\frac{C_t}{C_{t-n}} \times 100$, William's %R: $\frac{H_n - C_t}{H_n - L_n} \times 100$, A/D Oscillator: $\frac{H_t - C_{t-1}}{H_t - L_t} \times 100$, Disparity 5: $\frac{C_t}{MA_5} \times 100$, dan Moving Average: $\frac{\sum_{i=1}^n C_{t-i}}{n}$.
3. Gunakan data *training* untuk mendapatkan *hyperplane* terbaik.
4. Klasifikasikan data uji dengan menggunakan *hyperplane* tersebut
5. Bandingkan dengan data sebenarnya, diperoleh akurasi 55,56%,

4. Simpulan

Dalam memprediksi trend pergerakan harga saham Indeks LQ45, metoda HMM memiliki akurasi 50,98% sedangkan metoda SVM memiliki akurasi 55,56%. Untuk lebih akurat, selain menggunakan data histori (teknikal indikator), disarankan juga menggunakan informasi-informasi yang beredar tentang perusahaan dan juga faktor makro ekonomi.

Daftar Pustaka

- [1] Baresa, Suzana, Bogdan S, Ivanovic Z., 2013., Strategy of Stock Valuation by Fundamental Analysis, *UTMS Journal of Economics* 4(1), pp. 45-51.

- [2] Cristianini, N. and Taylor, JS., 2000 *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*, Cambridge University Press.
- [3] Hasan, Rafiul and Nath, Baikunth, 2005., *Stock Market Forecasting Using Hidden Markov Model: A New Approach*: IEEE Computer Society.
- [4] Jondri, Rizal.A, dan Saepudin.D., 2011., Deteksi Kelainan Paru-paru Menggunakan Wafelet dan Support Vector Machine, *Prosiding Seminar Nasional Matematika 2011, Unand, Padang*.
- [5] Rabiner, L.R., 1989., A Tutorial on Hidden Markov Models and Select Applications in Speech Recognition, *Proceedings of IEEE*, vol 77., No. 2.,pp 257-286.
- [6] Venkates, CK., T Yagi M, Ganesh, 2012., Fundamental Analysis and Stock Returns: An Indian Evidence, *Global Advanced Research Journal of economics, Accounting and Finance Vol. 1(2)*, pp. 033-039.