

## PREDIKSI HARGA CRUDE PALM OIL MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DENGAN OPTIMASI PARAMETER MENGGUNAKAN ALGORITMA GENETIKA

Oleh : Adiyanto

### ABSTRAK

Crude Palm Oil (CPO) merupakan minyak yang bersal dari buah kelapa sawit yang saat ini masih menjadi komoditi andalan dari Indonesia. Indonesia merupakan negara pengekspor CPO terbesar di dunia walapun saat ini telah tersaingi oleh Malaysia. Harga CPO Indonsia saat ini masih tergantung dari bursa forward Rotterdam. Prediksi harga CPO sangat diperlukan khususnya bagi para pembuat kebijakan, investor maupun perusahaan untuk menentukan patokan harga CPO, karena perkembangan harga CPO tergantung oleh beberapa faktor seperti iklim dan cuaca, harga minyak mentah dan komoditi subsidi. Metode *Support Vector machine* telah banyak digunakan untuk melakukan prediksi seperti prediksi  $PM_{10}$ , prediksi krisis keuangan dan prediksi kecepatan angin. Kelebihan Metode Support Vector Machine adalah dapat mengatasi masalah over fitting dan tidak memerlukan data yang terlalu banyak, serta memiliki tingkat akurasi yang tinggi. Kemampuan dari *Support Vecot Machine* yang memberikan dorongan kepada beberapa peneliti untuk mengembangkan penelitian lebih lanjut. Salah satu pengembangan lanjutan adalah memadukan *Support Vector Machine* dengan Algoritma Genetika yang digunakan untuk melakukan optimasi parameter. Algoritma Genetika digunakan untuk melakukan optimasi parameter yang menjadi kelemahan pada metode *Support Vector Machine*. Pada penelitian ini, eksperimen dengan menggunakan 3 jenis kernel yang pertama tipe kernel *dot* dengan RMSE sebesar 0,1289, akurasi sebesar 84,6%, yang kedua tie kernel *radial* dengan RMSE sebesar 0,0001, akurasi sebesar 98,5% dan yang ketiga tipe kernel *Polynomial* dengan RMSE sebesar 0,131 akurasi sebesar 83,3%.

Kata kunci : Prediksi, Harga, *Crude Palm Oil*, *Support Vector Machine*, Algoritma Genetika

### **Latar Belakang**

Indonesia merupakan negara produsen dan pengekspor *Crude Palm Oil* (CPO) terbesar di dunia (Syamsulbahri, 1996, p. 10). Meskipun saat ini kedudukannya sudah tergeser Negara Malaysia, namun demikian saat ini Indonesia masih mampu menyuplai kebutuhan minyak kelapa sawit dunia sekitar 5 juta ton pertahun. (Hadianto, 2003, p. 12).

Tingkat konsumsi CPO di dalam negeri sebesar 2.025 ribu ton, menunjukkan bahwa di dalam negeri sendiri pangsa

pasar CPO mendominasi sebagai bahan baku minyak goreng sebesar 82,22%, (Deperindo, 2007, p. 5). Kondisi tersebut merupakan pasar yang sangat berpotensi untuk lebih mengembangkan usaha kelapa sawit.

Peranan CPO dalam perdagangan minyak nabati dunia diperkirakan meningkat terus, di mana pertumbuhannya yang mencapai 5,4% per tahun melampaui perkembangan volume perdagangan jenis-jenis minyak nabati lainnya. Peningkatan tersebut di

atas cukup beralasan karena meningkatnya kecenderungan konsumsi CPO di beberapa Negara beberapa tahun terakhir ini.

Di beberapa Negara Eropa Barat kenaikan konsumsi CPO mencapai 9,6% pertahun dalam 5 tahun terakhir. Kenaikan tersebut dipicu oleh kebijakan pemerintah di Negara-negara Eropa Barat untuk mengganti penggunaan minyak lemak menjadi minyak nabati yang lebih sehat seperti CPO dan minyak nabati yang lainnya. Selain kebijakan tersebut kenaikan kebutuhan minyak kelapa sawit juga dipengaruhi oleh kebijakan beberapa negara yang berusaha untuk mengganti konsumsi minyak fosil menjadi CPO sebagai bahan bakar biodisel industri (Departemen Pertanian RI, 2011).

Penelitian untuk prediksi rentet waktu telah banyak dilakukan oleh peneliti lain dengan menggunakan beberapa metode diantaranya adalah metode *Support Vector Regression* (Arampongsanuwat & Meesad, 2011), *Support Vector Machine With Optimal Choice of Kernel Function Parameters* (Min & Lee, 2005), dan *Support Vector Machine* (Mohandes, Halawani, Rehman, & Hussain, 2004).

Metode lain yang digunakan untuk melakukan prediksi rentet waktu adalah Artificial Neural Network. Metode

Artificial Neural Network memiliki beberapa kelebihan diantaranya adalah kemampuannya melakukan pemodelan untuk pengenalan pola (Gupta, Madan, Jin, & Homma, 2003). Kelebihan lainnya adalah kemampuannya untuk melakukan pembelajaran berdasarkan data yang digunakan untuk pelatihan, dapat melakukan self organization atau melakukan representasi dari informasi yang diterimanya dan memiliki real time operation dalam pengertian Artificial Neural Network dapat melakukan perhitungan secara parallel serta memiliki fault toleran yang tinggi (Fausett, 1994). Selain memiliki kelebihan neural network juga memiliki kekurangan diantaranya membutuhkan data yang sangat besar untuk pelatihan dan memiliki konvergensi yang lambat.

*Support Vector Machine* (SVM) adalah sebuah algoritma yang menggunakan nonlinier mapping untuk mengubah data pelatihan asli menjadi data berdimensi tinggi (Sreelakshmi & Kumar, 2008, p. 1829). SVM pada awalnya dikembangkan untuk memecahkan masalah klasifikasi, kemudian kemampuannya telah dikembangkan untuk mengatasi regresi dan pengelompokan problem (Shi, Tian, Kou, Peng, & Lia, 2011, p. 3). Pada masa sekarang, SVM dengan linier atau nonlinier kernel telah menjadikan satu

dari kemampuan algoritma pembelajaran untuk klasifikasi sebaik untuk regresi dengan dua kemampuan utama pada data mining dengan menggunakan kernel mapping (Sreelakshmi, 2008, p. 1828).

Kemampuan lain yang dimiliki oleh metode SVM adalah memiliki akurasi yang tinggi dan tingkat kesalahan yang relative kecil, kemampuan untuk model kompleks nonlinear decision boundaries, lebih mudah untuk mengatasi over fitting, tidak membutuhkan data yang terlalu besar dan dapat digunakan untuk melakukan prediksi (Sreelakshmi & Kumar, 2008, p. 1829). Selain memiliki banyak kemampuan yang telah disebutkan di atas metode SVM juga memiliki kelemahan yaitu sulitnya menentukan parameter  $C$ ,  $\sigma$  dan  $\epsilon$  (Jian, Xibing, & Xiuzhi, 2012, p. 631).

Algoritma Genetika adalah algoritma yang berusaha menerapkan pemahaman mengenai evolusi alamiah pada tugas-tugas pemecahan masalah (*problem solving*) (Sumathi, Sivanandam, Lei, Chen, & Nguang, 2006, p. 402).

Pendekatan yang diambil oleh algoritma ini adalah dengan menggabungkan secara acak berbagai pilihan solusi terbaik di dalam suatu kumpulan untuk mendapatkan generasi solusi terbaik berikutnya yaitu pada suatu kondisi yang memaksimalkan kecocokannya

atau lazim disebut *fitness*. Generasi ini akan merepresentasikan perbaikan-perbaikan pada populasi awalnya.

Dengan melakukan proses ini secara berulang, algoritma ini diharapkan dapat mensimulasikan proses evolusioner.

Pada akhirnya, akan didapatkan solusi-solusi yang paling tepat bagi permasalahan yang dihadapi.

Algoritma Genetika memiliki beberapa kelebihan diantaranya adalah dapat digunakan untuk variabel diskrit maupun kontinyu, dapat digunakan untuk variabel yang besar, hasil akhir berupa beberapa variabel yang optimum, tidak hanya satu penyelesaian saja, optimasi dilakukan dengan mengkodekan variabel dan dapat digunakan pada data numerik, data eksperimental, atau fungsi analitik (Haupt & Haupt, 2004).

Dari uraian tersebut di atas, maka dalam penelitian ini akan digunakan metode *Support Vector Machine* yang dipadu dengan Algoritma Genetika yang akan digunakan untuk melakukan optimasi parameter *Support Vector Machine*.

### ***Support Vector Machine Berbasis Algoritma Genetika***

#### **Support Vector Regression**

Konsep dasar dari *Support Vector Regression* adalah untuk memetakan persamaan nonlinier dari data asli  $x$  menjadi sebuah space fitur berdimensi

tinggi (Wu et al., 2007 : 50), dan untuk menyelesaikan permasalahan regresi linear pada space fitur. Sebuah fungsi estimasi penaksiran regresi dapat ditentukan dengan memberikan data set  $T = \{(x_i, y_i)\}_i^l$  dimana  $x_i$  menyatakan input vector,  $y_i$  menyatakan koresponden nilai output dan  $i$  menyatakan jumlah dari data, fungsi regresi SVM adalah:

$$f(x) = w \cdot \varphi(x) + b$$

$\varphi(x)$  = space fitur berdimensi tinggi

$w$  = vector pembobot

$b$  = bias.

Pada persamaan 2.3 koefisien  $w$  dan  $b$  diestimasi dengan melakukan minimalisasi mengikuti fungsi resiko regular

$$R(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{l} \sum_{i=1}^l L_{\varepsilon}(y_i, f(x_i))$$

$C$  = ukuran fungsi nilai resiko empirik

$\frac{1}{2} \|w\|^2$  = syarat regulasi

$L_{\varepsilon}(y_i, f(x_i))$  =  $\varepsilon$ -insensitive loss function.

$\varepsilon$ -insensitive loss function dapat didefinisikan dengan persamaan sebagai berikut:

$$L_{\varepsilon}(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon & |y_i - f(x_i)| \geq \varepsilon \\ 0 & |y_i - f(x_i)| < \varepsilon \end{cases}$$

Pada persamaan 2.5 loss sama dengan nol jika kesalahan dari nilai prediksi

lebih kecil dari  $\varepsilon$ , selain itu loss melebihi  $\varepsilon$ .

Dua variabel positif  $\xi$  dan  $\xi^*$  dimasukan untuk merepresentasikan tujuan dari nilai actual menjadi nilai batasan koresponden dari  $\varepsilon$ -tube. Kemudian persamaan 2.4 ditransformasikan mengikuti bentuk

Minimize

$$\phi(w, \xi, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi + \xi^*)$$

Subject to

$$\begin{cases} y_i - [w \cdot \varphi(x)] - b \leq \varepsilon + \xi_i & \xi_i \geq 0 \\ [w \cdot \varphi(x)] + b - y_i \leq \varepsilon + \xi_i^* & \xi_i^* \geq 0 \end{cases}$$

Permasalahan optimasi diselesaikan dengan menggunakan bentuk

Lagrangian berikut:

Minimize

$$-\frac{1}{2} \sum_{j=1}^n (\alpha_j - \alpha_j^*) (\alpha_j - \alpha_j^*) (x_j, x_j) - \varepsilon \sum_{i=1}^n (\alpha_i - \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*)$$

Subject to

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \text{ and } 0 \leq \alpha_i, \alpha_i^* \leq C$$

Where

$\alpha_i - \alpha_i^*$  = lagrangian multiplier

$K(x_i, x_j)$  = kernel function

Nilai dari fungsi kernel sama dengan produk dari  $\varphi(x_i)$  dan  $\varphi(x_j)$  yang dibuat oleh pemetaan dari 2 vektor  $x_i$  dan  $x_j$  ke dalam space fitur berdimensi tinggi yaitu  $K(x_i, x_j) = \varphi(x_i) \varphi(x_j)$ .

Dengan penjumlahan Lagrange multiplier  $\alpha_i$  dan  $\alpha_j^*$  dioptimalkan menjadi vektor pembobot dari regresi *hyperplane* dengan mengikuti bentuk persamaan berikut:

$$w^* = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x_j)$$

Oleh karena itu, fungsi regresinya adalah

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x_j) + b$$

### Algoritma Genetika

Algoritma Genetika adalah suatu teknik optimasi yang didasarkan pada prinsip genetika dan seleksi alam (Shukla et al., 2010, p. 149). Algoritma Genetika merupakan metode pencarian yang disesuaikan dengan proses generika dari organisme biologi yang berdasar pada teori evolusi Charles Darwin (Melanie, 1999, p. 3). Algoritma Genetika terinspirasi dari mekanisme seleksi alam, dimana individu yang lebih kuat kemungkinan akan menjadi pemenang dalam lingkungan yang kompetitif dan solusi yang optimal dapat diperoleh dan diwakilkan oleh pemenang akhir dari permainan genetika (Haupt & Haupt, 2004, p. 22). Pada Algoritma Genetika tersedia solusi yang diterapkan pada sebuah populasi individu yang masing-masing mewakili solusi yang mungkin. Setiap solusi yang mungkin disebut dengan kromosom.

Secara umum struktur dari suatu algoritma genetik dapat mendefinisikan dengan langkah-langkah sebagai berikut:

1. *Membangkitkan populasi awal*  
Populasi awal ini dibangkitkan secara random sehingga didapatkan solusi awal. Populasi itu sendiri terdiri atas sejumlah kromosom yang merepresentasikan solusi yang diinginkan.
2. *Membentuk generasi baru*  
Untuk membentuk generasi baru, digunakan operator reproduksi seleksi, crossover dan mutasi. Proses ini dilakukan berulang-ulang sehingga didapatkan jumlah kromosom yang cukup untuk membentuk generasi baru dimana generasi baru ini merupakan representasi dari solusi baru. Generasi baru ini dikenal dengan istilah anak (*offspring*).
3. *Evaluasi solusi*  
Pada tiap generasi, kromosom akan melalui proses evaluasi dengan menggunakan alat ukur yang dinamakan *fitness*. Nilai *fitness* suatu kromosom menggambarkan kualitas kromosom dalam populasi tersebut. Proses ini akan mengevaluasi setiap populasi dengan menghitung nilai *fitness* setiap kromosom dan

mengevaluasinya sampai terpenuhi kriteria berhenti. Bila kriteria berhenti belum terpenuhi maka akan dibentuk lagi generasi baru dengan mengulangi langkah 2. Beberapa kriteria berhenti sering digunakan antara lain: berhenti pada generasi tertentu, berhenti setelah dalam beberapa generasi berturut-turut didapatkan nilai fitness tertinggi tidak berubah, berhenti dalam n generasi tidak didapatkan nilai fitness yang lebih tinggi.

Fungsi fitness tersebut sebagai berikut:

$$Fitness = \frac{1}{1 + penalti}$$

dimana

$$penalti = \sum Bp \sum Np$$

Dari persamaan diatas nilai fitness ditentukan oleh nilai penalti. Penalti tersebut menunjukkan jumlah pelanggaran kendala pada suatu kromosom. Semakin tinggi nilai fitness akan semakin besar kemungkinan kromosom tersebut terpilih ke generasi berikutnya. Jadi nilai penalti berbanding terbalik dengan nilai fitness, semakin kecil nilai penalti (jumlah pelanggaran) semakin besar nilai fitnessnya.

Fungsi fitness dapat didiperoleh dengan persamaan di bawah ini

$$Fitness = \frac{1}{1 + \sum Bp \sum Np} \quad 2.9$$

Keterangan :

$Bp$  = Bobot Pelanggaran

$Np$  = Indikator Pelanggaran

### Metode Evaluasi

Indikasi keberhasilan dari suatu prediksi adalah apakah hasil percobaan yang dilakukan telah akurat dan apakah model prediksi yang digunakan sudah optimal (Shukla et al., 2010, p. 58).

Upaya untuk mendapatkan jawaban atas pertanyaan-pertanyaan tersebut maka diperlukan sebuah metode pengukuran hasil dari suatu prediksi dalam bentuk perhitungan kesalahan.

Perhitungan kesalahan merupakan pengukuran untuk mengetahui apakah bila Algoritma Genetika diterapkan untuk pemilihan parameter pada metode *Support Vector Machine* dapat belajar dengan baik. Kesalahan keluaran hasil pada penerapan Algoritma Genetika apabila digunakan untuk pemilihan parameter *Support Vector Machine* dapat digunakan beberapa metode pengukuran kesalahan yaitu:

#### 1. Sum Square Error (SSE)

Perhitungannya adalah:

1. Hitung keluaran GA-*Support Vector Machine* untuk masukan pertama

2. Hitung selisih antara nilai keluaran GA-Support Vector Machine dan nilai target
3. Kuadratkan setiap keluaran kemudian hitung seluruhnya.

$$SSE = \sum_{i=1}^n (D_{ij} - f_j(x_i))$$

2. Root Mean Square Error Perhitungannya adalah:
  1. Hitung SSE
  2. Hasilnya dibagi dengan perkalian antara banyaknya data pada pelatihan dan banyaknya keluaran, kemudian diakarkan.

$$RMSE = \sqrt{\frac{SSE}{N * K}}$$

N = Banyaknya data pada pelatihan  
 K = Banyaknya Keluaran  
 Model yang telah optimal akan digunakan untuk simulasi terhadap data yang telah disiapkan untuk pengujian. Kinerja dari suatu model *support vector machine* setelah dilakukan suatu pelatihan dapat diukur dengan melihat *error* hasil pelatihan, validasi dan *testing* terhadap sekumpulan data *input* baru. Salah satu cara yang dapat digunakan untuk evaluasi ini adalah menggunakan analisis regresi terhadap respon jaringan dan target yang diharapkan.

**Optimsi Parameter SVM Dengan Menggunakan GA**

Pada penelitian ini motode yang diusulkan adalah dengan menggunakan

*Support Vector Machine* dengan menerapkan Algoritma Genetika untuk melakukan pemilihan parameter *Support Vector Machine*. Langkah-langkah pada Algoritma Genetika untuk penyeleksian parameter pada *Support Vector Machine* mengikuti algoritma berikut:

2.10

- Mulai
- Langkah 1 : Inisialisasi Parameter Kernel C dan ε.
  - Langkah 2 : Bangkitkan populasi dari kromosom dibangkitkan secara acak. Ukuran populasi di set ke 50.
  - Langkah 3 : Training SVM
  - Langkah 4 : Evaluasi Fitnes. Pada tahap ini fitness dari setiap kromosom dievaluasi, dengan menggunakan persamaan berikut:

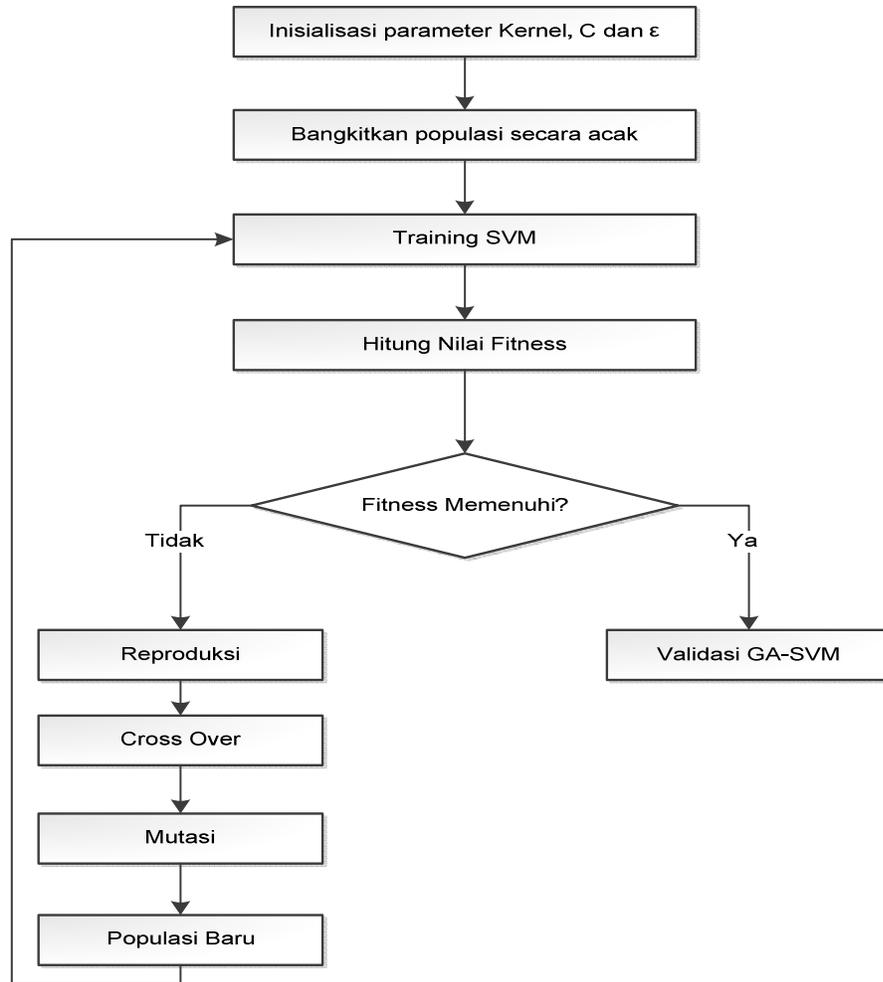
$$Fitness Function = \frac{1}{-\sigma^2 N} \sum_{i=1}^N (d_i - y_i)^2 \tag{2.11}$$

$$Where \sigma^2 = \frac{1}{N} \sum_{i=1}^N (d_i - \bar{d}_i)^2$$

N= total jumlah data set,  $\bar{d}_i$ = rata-rata nilai aktual,  $d_i$  = nilai aktual dan  $y_i$  = nilai prediksi

- Langkah 5 : Cek Nilai fitness, jika kondisi

- terpenuhi berhenti, selain itu
- Langkah 6 : Lakukan Reproduksi. Pada tahap ini populasi baru dibuat dengan perulangan mengikuti langkah-langkah sebelum populasi baru selesai
- Langkah 7 : Lakukan *crossover*. Dengan probabilitas *crossover*, *crossover* dari induk dibuat untuk membentuk offspring's (anak). Pada cross over, kromosom dipasangkan secara random. Prinsip *crossover* digunakan dan prosedur offspring mengikuti persamaan 3.4
- $$\mathbf{offspring = parent1 \pm a(parent2 - parent1) \quad 3.4}$$
- Langkah 8 : Lakukan mutasi. Setelah operasi *crossover* berhasil, string sebagai subyek untuk operasi mutasi, hal ini untuk mencegah
- runtuhnya seluruh solusi dari populasi menjadi local optimum dari penyelesaian masalah. Variabel dalam string yang akan bermutasi dipilih secara acak. Probabilitas *crossover* dan mutasi masing-masing diatur ke 0,8 dan 0,05.
- Langkah 9 : Populasi baru terbentuk, ulangi langkah 3.
- Selesai



Gambar 2.1 Metode Yang Diusulkan

**Eksperimen dan Pengujian Model**

Tahap eksperimen dilakukan untuk menghasilkan model prediksi dengan menggunakan langkah-langkah sebagai berikut:

1. Inisialisasi parameter  $C$ ,  $\gamma$  dan  $\epsilon$   
 Penentuan parameter akan berpengaruh terhadap kinerja sistem, range  $C$  yang ditetapkan pada GA-SVM adalah 60-500, range parameter  $\gamma$  dengan jarak antara 0.01 – 1.00, sedangkan range  $\epsilon$  yang ditetapkan adalah 0 – 1.0. Jumlah

kromosom dalam satu populasi adalah 10, peluang pindah silang adalah 0,8, dan peluang mutasi adalah 0,05. Sampel yang digunakan dalam penelitian ini adalah 50.

2. Pemilihan Kernel  
 Pemilihan kernel dilakukan untuk menentukan hasil dari kinerja sistem. Tipe kernel akan berpengaruh terhadap tingkat akurasi dan tingkat root mean square error hasil keluaran dari sistem. Tipe kernel yang digunakan dalam

penelitian ini adalah tipe polynomial, dot dan radial. Default tipe kernel pada *Rapidminer* adalah tipe dot.

### 3. Menentukan Max Iteration

Parameter ini akan berpengaruh terhadap kinerja sistem dalam proses pembelajaran terhadap proses prediksi. Proses akan berhenti setelah proses melakukan pengulangan sampai dengan batas maksimal yang ditentukan.

#### ***Hasil Eksperimen dan Pengujian Model/Metode***

Hasil dalam penelitian dilakukan dalam dua metode yaitu metode perhitungan

dengan menggunakan perhitungan Metode *Support Vector Machine* tanpa dikombinasi dengan Algoritma Genetika dan *Support Vector Machine* yang dikombinasi dengan Algoritma Genetika.

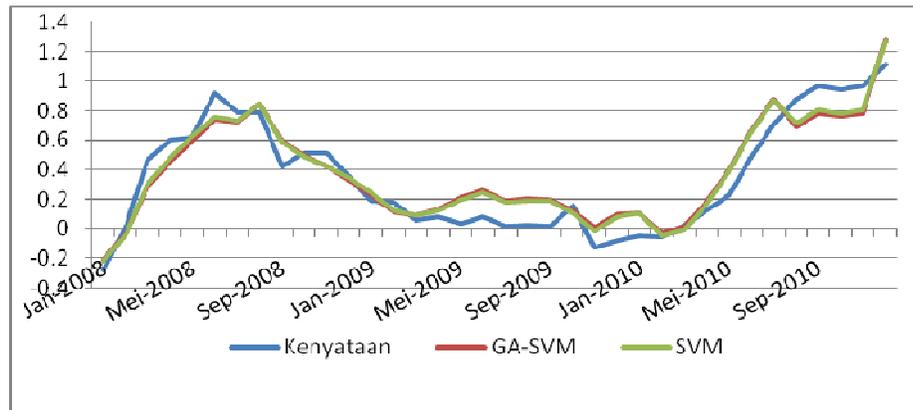
#### **Hasil Eksperimen dan Pengujian Model Menggunakan Tipe Kernel Polynomial**

Pengujian dilakukan dengan menggunakan tipe kernel polynomial pada *Support Vector machine* dengan dipadu oleh Algoritma Genetika untuk melakukan optimasi parameter *Support Vector Machine* dengan pola pengujian yang ditunjukkan dengan perubahan nilai parameter C, Gamma dan Epsilon.

Tabel 4.1 Pengujian Terbaik GA-SVM dan SVM Pada Tipe Kernel Polynomial

|                        | SVM      | GA-SVM   |
|------------------------|----------|----------|
| C                      | 239.7520 | 239.7708 |
| Gamma ( $\gamma$ )     | 0.4241   | 0.4190   |
| Epsilon ( $\epsilon$ ) | 0.0666   | 0.1625   |
| Akurasi                | 80.5%    | 83,3%    |
| RMSE                   | 0.065    | 0.131    |

Grafik perbandingan prediksi harga CPO pada SVM dan GA-SVM yang diterapkan pada data training ditunjukkan oleh gambar 3.1.



Gambar 4.1 Grafik Perbandingan Prediksi Harga GA-SVM dengan SVM

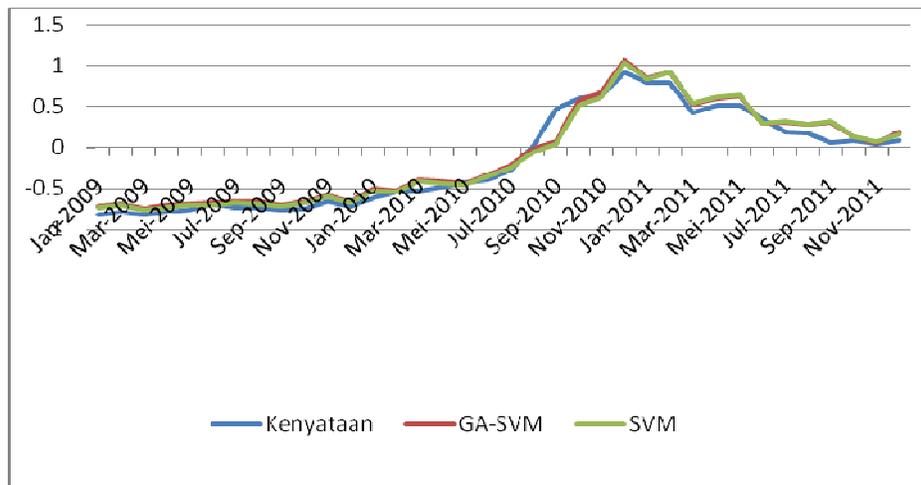
### Hasil Eksperimen dan Pengujian Model Menggunakan Tipe Kernel Dot

Pengujian ini dilakukan dengan menggunakan tipe kernel dot dan maximum iteration pada *Support Vector Machine* dengan menerapkan Algoritma Genetika. Eksperimen dengan menggunakan tipe kernel dot dilakukan sebanyak 20 kali dengan variasi perubahan input vektor. Indikator untuk mengetahui akurasi terbaik ditunjukkan oleh besarnya nilai akurasi dan RMSE untuk masing-masing pengujian pada data testing.

Tabel 4.2 Pengujian Terbaik SVM Dengan GA-SVM Pada Tipe Kernel Dot

|         | SVM      | GA-SVM   |
|---------|----------|----------|
| C       | 239.7589 | 239.7589 |
| Gamma   | 0.4253   | 0.4253   |
| Epsilon | 0.1683   | 0.1683   |
| Akurasi | 82,7%    | 84,6%    |
| RMSE    | 0,129    | 0.1289   |

Grafik perbandingan prediksi harga CPO pada metode SVM dengan metode GA-SVM yang diterapkan pada data training ditunjukkan oleh gambar 3.2.



Gambar 4.2 Grafik Perbandingan Prediksi Harga CPO Metode GA-SVM Dengan SVM

**Hasil Eksperimen dan Pengujian Model**

**Menggunakan Tipe Kernel Radial**

Pengujian ini dilakukan dengan menggunakan tipe kernel Radial dan maximum iteration pada *Support Vector Machine* dengan menerapkan Algoritma Genetika. Eksperimen dengan

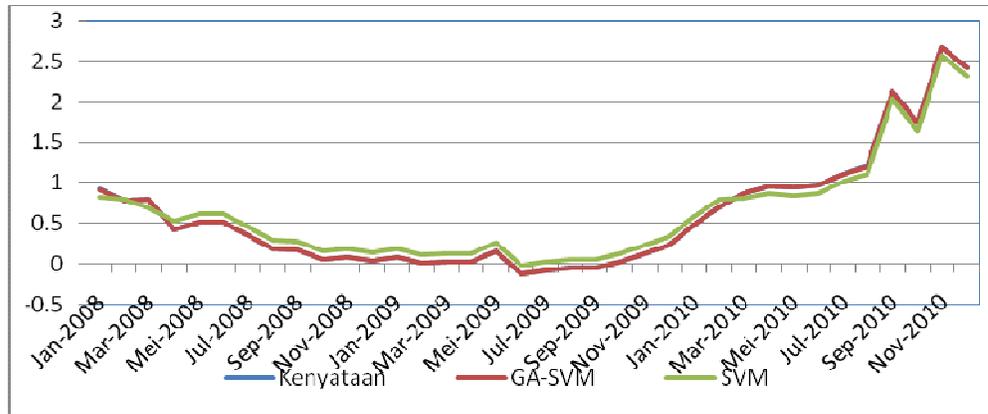
menggunakan tipe kernel radial dilakukan sebanyak 20 kali dengan variasi perubahan input vektor. Indikator untuk mengetahui akurasi terbaik ditunjukkan oleh besarnya nilai akurasi dan RMSE untuk masing-masing pengujian pada data testing.

Tabel 4.3 Hasil Pengujian Dengan *Support Vector Machine*

|         | SVM      | GA-SVM   |
|---------|----------|----------|
| C       | 239.7749 | 181.6168 |
| Gamma   | 0.3834   | 0.4394   |
| Epsilon | 0.0482   | 0.0001   |

|         |       |       |
|---------|-------|-------|
| Akurasi | 89,8% | 98.5% |
| RMSE    | 0,049 | 0.001 |

Grafik perbandingan prediksi harga CPO pada metode SVM dengan metode GA-SVM yang diterapkan pada data training ditunjukkan oleh gambar 4.3



Gambar 4.3 Grafik Perbandingan Prediksi Harga CPO Metode GA-SVM Dengan SVM

**Perbandingan Rata-Rata Nilai RMSE SVM dan GA-SVM**

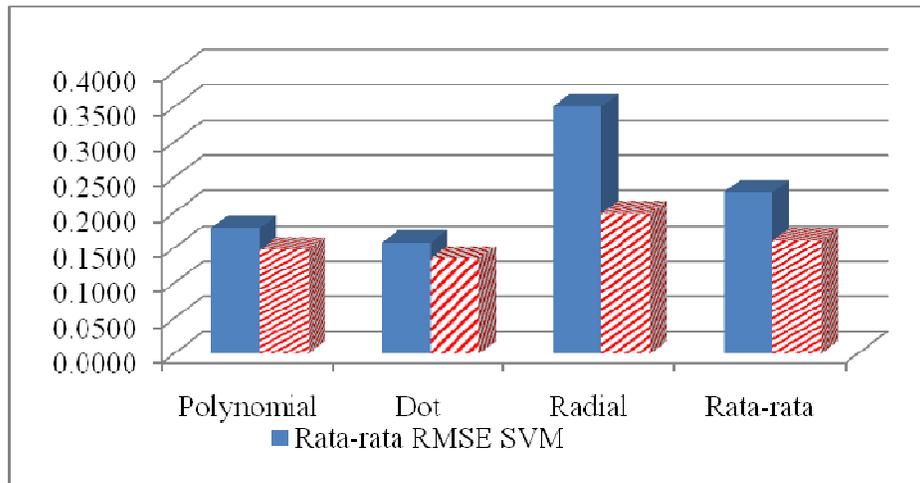
Berdasarkan hasil eksperimen dan analisis data dalam penelitian ini, maka dapat diperoleh perbedaan rata-rata nilai RMSE pada pengujian model SVM sebelum dan sesudah dilakukan optimasi parameter SVM dengan menggunakan Algoritma Genetika data yang

digunakan pada harga CPO. Perbandingan yang dihasilkan dapat memberikan informasi/gambaran tentang perbandingan rata-rata tingkat akurasi pada penerapan model tersebut. Tingkat perbandingan rata-rata nilai RMSE dapat dilihat pada tabel 4.4

Tabel 4.0. Perbandingan rata-rata Nilai RMSE Pada Tipe Kernel Radial

| Tipe Kernel | Rata-rata RMSE |        |
|-------------|----------------|--------|
|             | SVM            | GA-SVM |
| Polynomial  | 0.1782         | 0.1468 |
| Dot         | 0.1573         | 0.1348 |
| Radial      | 0.3505         | 0.1986 |
|             | 0.2286         | 0.1601 |

Dari data yang ditampilkan dalam Tabel 4.5, dapat ditampilkan grafik perbandingan rata-rata RMSE antara SVM dengan GA-SVM



Gambar 4.4 Grafik Perbandingan Rata-rata Nilai RMSE SVM Dengan GA-SVM

Berdasarkan tabel 4.4, menunjukkan bahwa pengujian penggunaan model GA-SVM pada tipe kernel Polynomial, dot dan radial untuk melakukan prediksi harga CPO memiliki rata-rata RMSE lebih kecil jika dibandingkan dengan model SVM, ini menunjukkan bahwa pada penggunaan tipe kernel polynomial, dot dan radial dapat meningkatkan peningkatan nilai akurasi prediksi harga CPO pada model SVM setelah dilakukan optimisasi parameter dengan Algoritma Genetika.

**Perbandingan Tingkat Akurasi Terbaik SVM dan GA-SVM**

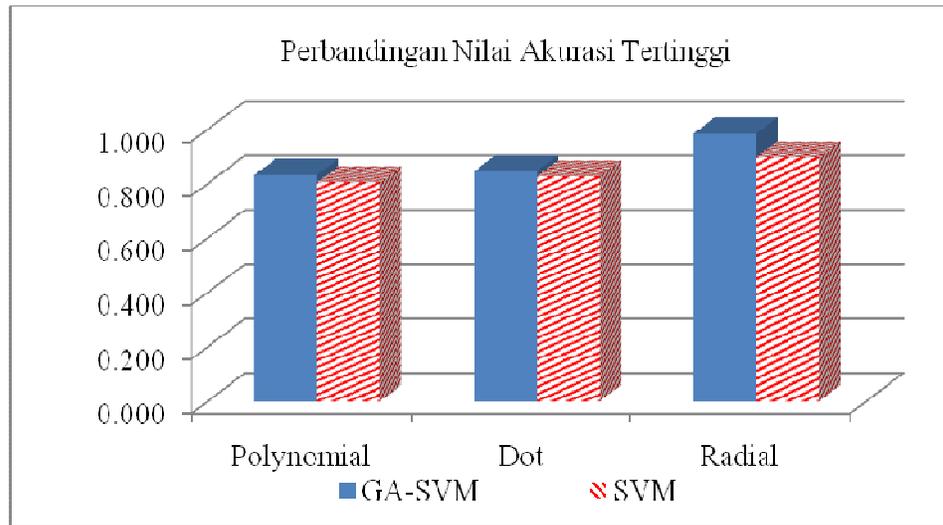
Berdasarkan hasil eksperimen dan analisis data dalam penelitian ini, maka dapat diperoleh perbandingan nilai akurasi tertinggi pada pengujian model SVM sebelum dan sesudah dilakukan optimisasi parameter dengan Algoritma Genetika pada data testing yang diterapkan pada prediksi harga CPO. Perbandingan ini dapat memberikan gambaran

tentang tingkat akurasi terbaik pada penerapan model tersebut. Perbandingan terbaik pada model tersebut dapat dilihat pada tabel 4.5  
Tabel 4.5 Perbandingan akurasi terbaik antara GA-SVM Dengan SVM

| Tipe   | GA-SVM    |         | SVM       |         |
|--------|-----------|---------|-----------|---------|
| Kernel | Parameter | Akurasi | Parameter | Akurasi |
|        |           |         |           |         |

|            | C       | $\gamma$ | $\epsilon$ |       | C       | $\gamma$ | $\epsilon$ |       |
|------------|---------|----------|------------|-------|---------|----------|------------|-------|
| Polynomial | 239.771 | 0.419    | 0.163      | 83.3% | 239.752 | 0.424    | 0.067      | 80.5% |
| Dot        | 239.759 | 0.415    | 0.168      | 84.6% | 239.759 | 0.415    | 0.168      | 82.7% |
| Radial     | 181.617 | 0.439    | 0.000      | 98.5% | 239.775 | 0.383    | 0.048      | 89.8% |

Perbandingan akurasi terbaik berdasarkan tabel 4.5 jika diilustrasikan dalam grafik maka akan tampak seperti gambar 4.5



Gambar 4.5 Grafik Perbandingan Terbaik Prediksi Harga CPO dengan Metode GA-SVM dan SVM

Mengacu tabel 4.5, menunjukkan bahwa penerapan model SVM sebelum dilakukan optimasi parameter nilai akurasi tertinggi terjadi pada tipe kernel polynomial dengan nilai  $C = 239.752$ ,  $\epsilon = 0.067$  dengan nilai akurasi sebesar 80.5%, nilai akurasi tertinggi pada tipe kernel dot terjadi pada nilai  $C = 239.759$ ,  $\epsilon = 0.168$  dan nilai akurasi tertinggi pada tipe kernel radial dengan  $C = 239.775$ ,  $\gamma = 0.383$ , dan  $\epsilon = 0.048$  dengan nilai akurasi sebesar 89,8%. Sedangkan penerapan SVM setelah dilakukan optimasi parameter dengan menggunakan GA nilai akurasi tertinggi pada tipe kernel polynomial terjadi pada nilai C

239.771,  $\gamma = 0.419$ , dan  $\epsilon = 0.163$  dengan tingkat akurasi sebesar 83,3%, untuk nilai akurasi tertinggi pada tipe kernel dot terjadi pada nilai  $C = 239.759$   $\gamma = 0.415$  dan  $\epsilon = 0.168$  dengan nilai akurasi sebesar 84.6% dan untuk nilai akurasi tertinggi pada tipe kernel radial terjadi pada nilai  $C = 181.617$   $\gamma = 0.439$  dan  $\epsilon = 0.000$  dengan nilai akurasi sebesar 98,5%.

Dalam penelitian ini ditemukan bahwa tingkat akurasi tertinggi pada metode SVM sebelum dilakukan optimasi parameter dengan Algoritma Genetika adalah terjadi pada tipe kernel radial dengan nilai akurasi sebesar 89,8%. Sedangkan metode SVM setelah

dilakukan optimasi parameter nilai akurasi tertinggi terjadi pada tipe kernel radial dengan nilai akurasi sebesar 98,5%.

### **Kesimpulan**

Pengujian penerapan algoritma genetika yang digunakan untuk melakukan optimasi parameter  $C$ ,  $\gamma$  dan  $\epsilon$  pada metode support vector machine dalam prediksi harga CPO yang dilakukan dengan menggunakan 60 eksperimen dengan menggunakan tiga tipe kernel yaitu kernel dot, radial, dan polynomial. Dari hasil pengujian ditemukan bahwa rata-rata RMSE dengan menggunakan metode SVM sebelum dioptimasi sebesar 0,2142 sedangkan rata-rata RMSE dengan menggunakan metode GA-SVM sebesar 0,1300, dari hasil tersebut berarti ada penurunan RMSE sebesar 0,0842. Sedangkan rata-rata akurasi dengan menggunakan metode SVM sebelum dioptimasi sebesar 68,1% sedangkan rata-rata akurasi pada metode GA-SVM sebesar 74,5%. Akurasi terbaik diperoleh dari tipe kernel radial dimana pada metode SVM sebesar 84,6% sedangkan pada metode GA-SVM mencapai 98,5%.

### **DAFTAR PUSTAKA**

- Arampongsanuwat, S., & Meesad, P. (2011). Prediction of PM 10 using Support Vector Regression. *International Conference on Information and Electronics Engineering*, 6, 120–124.
- Chen, B., & Chen, M. (2011). Applying particles swarm optimization for support vector machines on predicting company financial crisis. *International Conference on Business and Economic Research*, 1, 301–305.
- Dawson, C. W. (2009). *Project in computing and information system* (Second Edi.). Addison Wesley.
- Fausett, L. (1994). *Fundamental Of Neural Network Architecture, Algorithms, And Applications*. New Jersey: Prentice Hall.
- Gupta, Madan, M. ., Jin, L., & Homma, N. (2003). *Static And Dynamic Neural Networks From Fundamentals to Advanced Theory*. A John Wiley and Sons, Ltd.
- Haupt, R. L. ., & Haupt, S. E. (2004). *Practical Genetic Algorithms* (Second Edi.). New Jersey: A John Wiley and Sons, Ltd.
- Indonesia, D. P. (2007). *Gambaran Sekilas Industri Minyak Kelapa Sawit*. Jakarta: World Growth Palm Oil Development Campaign.
- Jian, Z., Xibing, L., & Xiuzhi, S. (2012). Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines q. *Elsevier*, 50(4), 629–644. doi:10.1016/j.ssci.2011.08.065
- Melanie, M. (1999). *An Introduction to Genetic Algorithms* (Five EDITION.). London: MIT Press.

- Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Elsevier*, 28(4), 603–614. doi:10.1016/j.eswa.2004.12.008
- Badan Promosi Dan Investasi Daerah. (2009). Profil Proyek Komoditi Unggulan Daerah Kalimantan Timur. *Badan Perijinan Dan Penanaman Modal Daerah*.
- Mohandes, M. A., Halawani, T. O., Rehman, S., & Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Elsevier*, 29, 939–947. doi:10.1016/j.renene.2003.11.009
- Petrova, N. V., & Wu, C. H. (2006). Prediction of catalytic residues using Support Vector Machine with selected protein sequence and structural properties. *BMC Bioinformatics*, 7(1), 312.
- Departemen Pertanian RI. (2011). *Manfaat Minyak Sawit bagi Perekonomian Indonesia. World Growth Palm Oil Development Campaign*. World Growth Palm Oil Development Campaign.
- Kementrian Perdagangan RI. (2010). Kontrak Berjangka. *Badan Pengawasan Perdagangan Berjangka Komoditi*.
- Shi, Y., Tian, Y., Kou, G., Peng, Y., & Lia, J. (2011). Optimization Based Data Mining Theory and Application. *Springer London*. doi:10.1007/978-0-85729-504-0
- Shukla, A., Tiwari, R., & Kala, R. (2010). *Real Life Application of Soft Computing*. CRC Press.
- Sreelakshmi, K., & Kumar, P., R. (2008). Short Term Wind Speed Predicting Using Support Vector Machine Model. *WSEAS Transactions*, 7(11), 1828–1837.
- Sumathi, S., Sivanandam, S. N., Lei, V., Chen, Z., & Nguang, S. K. (2006). *Introduction to Data Mining and its Applications Studies in Computational Intelligence (Volume 29)*. London: Springer.
- Vapnik, V. (1998). *The Nature Of Statistik Learning Theory*. John Wiley & Sons.
- Warta Penelitian Dan Pengembangan Pertanian. (2009). Membentuk Harga Referensi CPO Dunia di Indonesia, 16–18.
- Weise, T. (2009). *Global Optimization Algorithms – Theory and Application – (Second Edi.)*. it-weise.de (self-published). Retrieved from <http://www.it-weise.de/documents/metaWGOEB.html>
- Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., et al. (2007). *Top 10 algorithms in data mining (Vol. 14, pp. 1–37)*. CRC Press. doi:10.1007/s10115-007-0114-2