



Computer Based Information System Journal

ISSN (Print): 2337-8794 | E- ISSN : 2621-5292
web jurnal : <http://ejournal.upbatam.ac.id/index.php/cbis>



PREDIKSI HARGA CABAI MERAH MENGGUNAKAN *SUPPORT VECTOR REGRESSION*

Domi Sepri¹⁾, Ahmad Fauzi²⁾, Raju Wandira³⁾, Ozzy Secio Riza⁴⁾, Yeni Fitri Wahyuni⁵⁾, Hummaira Hutagaol⁶⁾

1) 2) 3) 4) 5) 6) UIN Imam Bonjol, Jl. Prof Yunus, Padang, Sumatera Barat, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: Juli 2020
Diterbitkan *Online*: September 2020

KATA KUNCI

Support Vector Regression, Harga Cabai Nasional, Time Series, Grid Search

KORESPONDENSI

E-mail:
domisepri@uinib.ac.id

A B S T R A C T

The aim of the research is to make a model to predict the national price of the red chili pepper using Support Vector Regression (SVR). Data used to make model are the price of red chili from January 2017 to December 2019. For finding the best parameter of the hyperplane, the research uses Grid Search Algorithm. The best parameter of the hyperplane is $C=1000$, $\epsilon = 5$, and $\gamma = 1$. The result shows the MAPE for data training is 4.07% and the MAPE for data testing is 9.11%.

I. Latar Belakang

Cabai bagi masyarakat Indonesia merupakan komoditas sayuran yang sangat penting karena merupakan bahan utama dalam pengolahan berbagai makanan. Pada masa-masa tertentu harga cabai yang tidak stabil memicu timbulnya inflasi perekonomian seperti saat memasuki masa musim hujan, kemarau atau saat perayaan hari besar keagamaan di Indonesia. Untuk mengantisipasi harga cabai di pasaran agar tetap dalam batas yang wajar, pendekatan yang perlu dilakukan oleh pemerintah yaitu salah satunya dengan melakukan prediksi harga pangan cabai.

Beberapa metode *stockistik* dan mesin pembelajaran telah dikembangkan untuk melakukan prediksi terhadap data yang bergantung pada waktu (*Time Series*). Salah satu algoritme untuk melakukan prediksi bergantung waktu yaitu dengan menggunakan *Support Vector Regression* (SVR). [1] melakukan penelitian dengan menggunakan SVR dengan *Kernel Anova* dan *Gaussian RBF* untuk memprediksi harga batubara. Data yang digunakan oleh [1] merupakan harga batu bara periode Januari 2009 sampai dengan Desember 2017. Hasil penelitian [1] menunjukkan SVR mampu melakukan prediksi terhadap harga batu bara. Penggunaan SVR juga banyak ditemukan

pada penelitian terkait dengan harga saham, di antaranya penelitian yang dilakukan oleh [2], [3], [4].

Tujuan penelitian ini yaitu membuat pemodelan untuk memprediksi harga cabai merah nasional dengan menggunakan algoritma SVR. Pemodelan prediksi ini dilakukan sebagai salah satu alternatif pedoman bagi pemerintah dan pihak terkait dalam membuat kebijakan yang berhubungan dengan harga cabai. Penelitian ini diharapkan dapat bermanfaat bagi masyarakat luas sehingga kestabilan harga cabai nasional di Indonesia dapat terwujud.

II. Kajian Literatur

Support Vector Machine (SVM) pertama kali diperkenalkan oleh Vapnik untuk menggeneralisasi data yang tidak terlihat dengan mengimpelentasikan algoritme pembelajaran [5] [6]. Cara kerja SVM yaitu dengan mencari maksimum margin dari *hyperplane* dan mencari *error* terkecil. Data yang paling dekat dengan maksimum margin *hyperplane* disebut *Support Vector*.

Secara umum persamaan SVM ditunjukkan oleh persamaan (1)

$$f(x) = w\varphi(x) + b \tag{1}$$

dengan $\varphi(x)$ adalah fungsi yang memetakan input x kedalam ruang *feature*, b adalah koefisien bias, dan W adalah faktor bobot. *Error* terkecil diperoleh Ketika

$$\min \frac{1}{2} \|w\|^2 \tag{2}$$

dengan *constraint*

$$y_i - \langle w, \varphi(x_i) \rangle - b \leq \varepsilon \tag{3}$$

$$\langle w, \varphi(x_i) \rangle + b - y_i \leq \varepsilon \tag{4}$$

Penambahan *slack* ξ dan ξ^* digunakan untuk melakukan optimasi terhadap *error* yang sudah ada. Sehingga persamaan (2) menjadi

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \tag{5}$$

dengan *constraint*

$$y_i - \langle w, \varphi(x_i) \rangle - b \leq \varepsilon + \xi_i \tag{6}$$

$$\langle w, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \tag{7}$$

$$\xi_i, \xi_i^* \geq 0 \tag{8}$$

Ada dua kelompok utama dari SVM, *Support Vector Classifier* (SVC) dan *Support Vector Regression* (SVR). SVC merupakan SVM untuk menyelesaikan permasalahan terkait data diskret sedangkan SVR merupakan SVM yang digunakan untuk menyelesaikan kasus regresi atau data continue. SVR diajukan pertama kali oleh Vapnik, Steven Golowich, dan Alex Smola pada tahun 1997 [7]. SVR pada pengimplentasiannya menggunakan *high dimensional feature space* atau disebut dengan Kernel. Tujuan dari Kernel yaitu mentransformasi data ke dalam ruang feature yang lebih tinggi. Persamaan SVR untuk kasus Non-Linear ditunjukkan oleh persamaan (9)

$$f(x) = \sum_i^n (\alpha_i + \alpha_i^*) K(x_i, x) + b \tag{9}$$

dengan α dan α^* adalah *Lagrange Multiplier*, $K(x_i, x)$ adalah fungsi Kernel.

Kernel pada SVR yang banyak digunakan dalam penelitian adalah *Radial Basis Function* (RBF) (persamaan (10)).

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2} \|x_i - x_j\|^2\right) \tag{10}$$

III. Metodologi

Penelitian ini menggunakan data harga cabai merah nasional periode Januari 2017 sampai dengan Desember 2019. Data dibagi menjadi dua kelompok, data latih dan data uji. Jumlah data latih sebanyak 80% dari data keseluruhan, sedangkan jumlah data uji sebanyak 20% dari data keseluruhan.

Pembuatan model dilakukan dengan memasukkan parameter *hyperplane* SVR. Penentuan parameter terbaik dari *hyperplane* dilakukan dengan menggunakan *Grid Search*

dengan Kernel RBF. Algoritme *Grid Search* pada penelitian ini menggunakan 10-Fold Cross validation dengan variasi C [1,100,1000], ϵ [0.01 0.05, 0.1, 0.5, 1, 5] dan γ [1,3,5].

Evaluasi hasil prediksi dilakukan dengan menggunakan *Mean Absolute Percentage Error* (MAPE). MAPE dihitung dengan membandingkan selisih antara nilai aktual pada waktu ke-t (Y_t) dengan nilai peramalan pada waktu ke-t (F_t), dengan n adalah banyak nya data.

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right|}{n} \quad (11)$$

Skala akurasi MAPE untuk prediksi ditunjukkan oleh Tabel 1

Table 1 Akurasi Prediksi [8]

MAPE	Akurasi Prediksi
<10 %	sangat baik
11%-20%	Baik
21%-50%	masih dalam batas wajar
>51%	tidak akurat

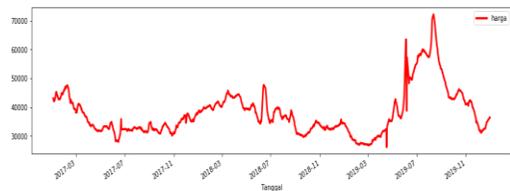
IV. Pembahasan

Tabel 2 di bawah menunjukkan deskripsi dari data yang digunakan. Harga cabai merah terendah periode Januari 2017 sampai dengan Desember 2019 adalah Rp. 26,100, sedangkan harga tertinggi cabai merah periode yang sama Rp. 72,200. Dari data tersebut diketahui bahwa harga cabai nasional tidak bersifat musiman (Gambar 1). Pada tahun 2019, harga cabai nasional relatif mengalami peningkatan yang signifikan dibandingkan tahun 2017 dan 2018 (Gambar 1).

Table 2 Deskripsi data

Jumlah	743
mean	38105.921938

std	8144.577922
min	26100
25%	32425
50%	35800
75%	41850
max	72200



Gambar 1 Data Harga Cabai Nasional Januari 2017 s.d Desember 2019

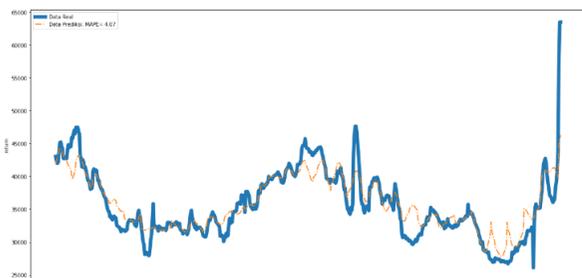
Berdasarkan hasil pencarian parameter *hyperplane* terbaik dengan menggunakan Algoritme *Grid Search*, diperoleh parameter *hyperplane* terbaik yaitu pada C=1000, $\epsilon = 5$ dan $\gamma = 1$ dengan nilai *mean test* sebesar 0.777577 (Table 3).). Ketika nilai C semakin besar, maka hal tersebut membuat margin akan semakin kecil, namun memiliki variance yang besar. Sebaliknya untuk nilai γ . Di dalam proses SVM sendiri dengan menggunakan kernel RBF, γ mengatur area dari support vector.

Parameter yang sudah diperoleh dengan menggunakan *Grid Search* kemudian dilanjutkan dengan melakukan pemetaan terhadap data latih dan data uji dan dievaluasi menggunakan MAPE. Pada pengujian dan evaluasi prediksi dengan menggunakan data latih diperoleh MAPE sebesar 4.07% (Gambar 2). Sedangkan evaluasi hasil prediksi data uji dengan data sebenarnya diperoleh MAPE sebesar 9.11% (Gambar 3). MAPE pada data uji lebih besar dibandingkan data latih menunjukkan tidak adanya *overfitting* dalam proses pembuatan model. Selain itu, baik MAPE data latih dan data uji masih berada pada daerah akurasi prediksi sangat baik.

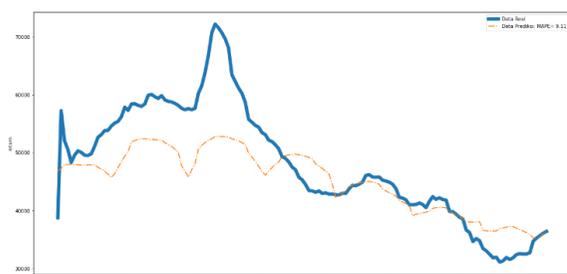
Table 3 Hasil pencarian parameter terbaik dengan menggunakan Grid search untuk Kernel RBF dan $C_v=10$

Table 3 Hasil pencarian parameter terbaik dengan menggunakan Grid search untuk Kernel RBF dan $C_v=10$				C	ϵ	γ	mean_test
				100	0,1	3	0,111442
				100	0,5	3	0,11144
				100	1	3	0,111439
1000	5	1	0,777577	100	5	3	0,111406
1000	1	1	0,777569	100	0,01	5	0,050132
1000	0,5	1	0,777564	100	0,05	5	0,050132
1000	0,1	1	0,77756	100	0,1	5	0,050131
1000	0,05	1	0,77756	100	0,5	5	0,050126
1000	0,01	1	0,777559	100	1	5	0,05012
1000	5	3	0,72986	100	5	5	0,050069
1000	1	3	0,729835	1	5	1	-0,08001
1000	0,5	3	0,729831	1	0,01	1	-0,08005
1000	0,1	3	0,729828	1	0,1	1	-0,08005
1000	0,05	3	0,729827	1	0,05	1	-0,08005
1000	0,01	3	0,729827	1	0,5	1	-0,08005
1000	0,01	5	0,643628	1	1	1	-0,08005
1000	0,05	5	0,643628	1	5	3	-0,08248
1000	0,1	5	0,643627	1	1	3	-0,08254
1000	0,5	5	0,643623	1	0,5	3	-0,08254
1000	1	5	0,643618	1	0,05	3	-0,08254
1000	5	5	0,643581	1	0,01	3	-0,08254
100	5	1	0,241133	1	0,1	3	-0,08254
100	1	1	0,241121	1	5	5	-0,08327
100	0,5	1	0,241119	1	1	5	-0,08332
100	0,1	1	0,241118	1	0,5	5	-0,08333
100	0,05	1	0,241118	1	0,1	5	-0,08333
100	0,01	1	0,241118	1	0,05	5	-0,08333
100	0,01	3	0,111442	1	0,01	5	-0,08333
100	0,05	3	0,111442				

Daftar Pustaka



Gambar 2 Pemodelan Data Latih



Gambar 3 Pemodelan Data Uji

V. Kesimpulan

Hasil *tunning* parameter *hyperplane* dengan menggunakan algoritme Grid Search pada Kernel RBF, diperoleh parameter terbaik $C=1000$, $\epsilon = 5$ dan $\gamma = 1$. Hasil pemodelan dengan menggunakan SVR menunjukkan MAPE prediksi untuk data uji lebih besar dibandingkan MAPE prediksi data latih. MAPE data latih = 4.07% dan MAPE data uji= 9.11%.

Pemodelan selanjutnya diharapkan dapat memperkecil MAPE atau *error* data latih dan data uji. Pengecilan *error* dapat dilakukan dengan melakukan *ensemble learning* dengan beberapa algoritme ataupun dengan mentransformasi data ke dalam ruang *feature* yang lain.

[1] O. Bonita, L. Muflikhah and R. K. Dewi, "Prediksi Harga Batu Bara Menggunakan Support Vector Regression (SVR)," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 2, no. 12, pp. 6603-6609, 2018.

[2] R. Maharesi, "Penggunaan Support Vector Regression (SVR) Pada Prediksi Return Saham Syariah BEI," in Proceeding PESAT (Psikologi, Ekonomi, Sastra, Arsitektur & Teknik Sipil), Bandung, 2013.

[3] H. Yasin, A. Prahutama and T. W. Utami, "Prediksi Harga Saham Menggunakan Support Vector Regression Dengan Algoritma Grid Search," Media Statistika, vol. 7, no. 1, pp. 29-35, 2014.

[4] B. M. Henrique, V. A. Sobreiro and H. Kimura, "Stock Price Prediction Using Support Vector Regression On Daily And Up To The Minute Prices," The Journal of Finance and Data Science, vol. 4, pp. 183-201, 2018. [5] A. J. Smola and B. Scho" Lkopf, "A Tutorial on Support Vector Regression," Kluwer Academic Publishers, Netherlands, 2004.

[6] D. Basak, S. Pal and D. C. Patranabis, "Support Vector Regression," Neural Information Processing – Letters and Reviews, vol. 11, no. 10, pp. 2013-224, 2007.

[7] V. Vapnik, S. Golowich and A. Smola, Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing, Cambridge: MIT Press, 1997.

[8] K. D. Lawrance, R. K. Klimberg and S. M. Lawrance, Fundamentals of Forecasting Using Excel, America: Insdustrial Press Inc, 2009.