

ANALISIS SUPPORT VECTOR REGRESSION (SVR) DENGAN KERNEL RADIAL BASIS FUNCTION (RBF) UNTUK MEMPREDIKSI LAJU INFLASI DI INDONESIA

Isnaeni R¹, Sudarmin¹, Zulkifli Rais¹

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: Backpropagation
Neural Network;
Indikator Kesejahteraan
Rakyat;
Klasifikasi

Abstract:

Inflation is one indicator that affects the economic growth of a country. As a developing country, Indonesia has an unstable inflation rate every year. Therefore, it is necessary to predict the inflation rate in the future to be useful for formulating future economic policies. SVR is a Support Vector Machine (SVM) development for regression cases. In the SVR method, the RBF kernel is used as an aid in solving non-linear problems, the Min – Max Normalization method for data normalization, distribution of training data and testing data, selecting the best model with Grid Search Optimization, then forecasting using the model obtained with parameter = 0,1, C = 1, and = 3. The forecasting results obtained were evaluated by looking at the RMSE value, the test value obtained was RMSE of 0.0020, which means the model's ability to follow the data pattern well.

1. Pendahuluan

Prediksi adalah sebuah seni dan ilmu untuk meramalkan kejadian di masa mendatang. Peramalan dapat dilakukan dengan melibatkan data masa lalu (historis) kemudian memproyeksikannya ke masa depan dengan suatu bentuk model matematis (Maulana dkk., 2019). Metode yang sering digunakan dalam pengklasifikasian ataupun peramalan adalah Support Vector Machine, Backpropagation, Radial basis Function, K- Nearest Network, Analisis Diskriminan, Simple Logistic Classifier (SLC), Fuzzy, dan masih banyak lagi.

Metode Support Vector Machine (SVM) banyak digunakan dalam penelitian dikarenakan metode ini mampu menyesuaikan dengan jenis data yang digunakan serta minimnya kesalahan error. Hasil menunjukkan efisiensi dan akurasi yang baik. Metode SVM banyak digunakan untuk melakukan penelitian klasifikasi otomatis (Widiastuti dkk., 2017). Sedangkan untuk menyelesaikan masalah regresi, diterapkan metode Support Vector Regression (SVR). Algoritma SVR adalah teori yang diadaptasi dari teori SVM untuk kasus regresi yang menghasilkan keluaran berupa bilangan riil. Konsep algoritma SVR dapat menghasilkan nilai peramalan yang bagus karena SVR mempunyai kemampuan menyelesaikan masalah overfitting. Overfitting adalah perilaku data saat data testing atau training menghasilkan akurasi prediksi hampir sempurna (Furi dkk., 2015).

Untuk penelitian ini, akan digunakan data inflasi berdasarkan indeks harga konsumen (IHK). Inflasi dapat diartikan sebagai kenaikan harga barang dan jasa secara umum dan terus menerus dalam jangka waktu tertentu. Untuk mengukur tingkat inflasi, indikator yang banyak digunakan adalah IHK. Perubahan IHK menunjukkan perubahan harga

* Corresponding author.

E-mail address: zulkifli.rais89@unm.ac.id.



barang dan jasa yang dikonsumsi masyarakat dari waktu ke waktu. Penetapan barang dan jasa dalam perhitungan IHK didasarkan pada Survei Biaya Hidup (SBH) yang dilakukan secara berkala oleh BPS.

Kenaikan harga dari satu atau dua barang saja tidak dapat disebut inflasi kecuali bila kenaikan itu meluas (atau mengakibatkan kenaikan harga) pada barang lainnya (Bank Indonesia, 2021). Karena data inflasi merupakan jenis data non – linier, maka peneliti memilih menggunakan metode SVR untuk mengolah data tersebut. Metode SVR diterapkan dengan kernel Radial Basis Function (RBF), dengan demikian metode diatas sangat cocok untuk menangani masalah hubungan non – linier sesuai dengan data yang akan digunakan (Caraka, 2017)

Pada jurnal penelitian sebelumnya yang menggunakan metode SVR yaitu, “Implementasi Support Vector Regression pada Prediksi Inflasi Indeks Harga Konsumen,” “Forecasting Inflation Rate Using Support Vector Regression (SVR) based Weight Attribute Particle Swarm Optimization (WAPSO),” dan “Penerapan Support Vector Regression (SVR) untuk Peramalan Inflasi Bulanan Nasional,”. Metode SVR yang diterapkan pada jurnal penelitian diatas memberikan kesimpulan bahwa metode SVR sudah cukup baik dalam hasil prediksinya. Kesalahan yang dihasilkan sudah cukup minim dengan bantuan kernel yang berbeda – beda.

2. Tinjauan Pustaka

SVR merupakan pengembangan SVM untuk kasus regresi. Tujuan dari SVR adalah untuk menemukan sebuah fungsi sebagai suatu hyperplane (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah error dan membuatnya sekecil mungkin (Smola & Sch, 2004). Misalkan ada data training, dimana merupakan vektor input dan output skalar dan adalah banyaknya data training. Dengan SVR, ingin ditentukan suatu fungsi yang mempunyai deviasi paling besar dari target aktual, untuk semua data training. Jika nilai sama dengan 0 maka diperoleh suatu persamaan regresi yang sempurna (Amnda, 2014). Support Vector (SV) adalah sebuah metode statistik yang memiliki konsep seperti jaringan saraf, dan fungsi radial basis sebagai sebuah jaringan kasus khusus. Dalam kasus RBF, algoritma SV otomatis menentukan pusat, bobot dan threshold seperti untuk meminimalkan batas atas kesalahan pengtestingan yang diharapkan. Namun demikian, hasil yang sangat baik telah diperoleh dengan menggunakan SVR dalam waktu masalah seri SVR (Caraka, 2017).

SVM adalah sebuah metode yang digunakan untuk kasus klasifikasi, namun prinsip metode tersebut dapat dikembangkan ke dalam regresi dan metode peramalan (Time Series). Misalkan terdapat l data training, (x_i, y_i) , $i = 1, \dots, l$ dengan data input $x = \{x_1, \dots, x_l\} \subseteq \mathcal{R}^N$ dan $y = \{y_1, \dots, y_l\} \subseteq \mathcal{R}$ dan l adalah banyaknya data training. Fungsi regresi dari metode SVR adalah sebagai berikut:

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

dimana :

- w : vektor pembobot
- $\phi(x)$: fungsi yang memetakan x dalam suatu dimensi
- b : bias

Dalam mendukung menyelesaikan permasalahan non-linier dengan algoritme SVR, maka digunakanlah fungsi kernel. Kelebihan dari penggunaan fungsi kernel ini yaitu mampu menyelesaikan masalah non - linier yang berdimensi lebih tinggi tanpa perlunya perhitungan pemetaan eksplisit (Furi dkk., 2015)

Ada beberapa fungsi kernel yang sering digunakan dalam literature SVM diantaranya kernel linear, kernel polynomial, kernel Radial Basic Function (RBF) atau biasa disebut dengan kernel Gaussian dan kernel sigmoid. Rumus fungsi kernel tersebut terdapat pada Tabel 1 (Septiningrum dkk., 2015)

Tabel 1. Jenis – jenis Kernel

Nama kernel	Fungsi
Gaussian RBF	$K(x_i, x_j) = \exp(-\gamma(x_i, x_j)^2)$



<https://doi.org/10.35580/variensiunm6>

Linear	$K(x_i, x_j) = x_i, x_j$
Polynomial	$K(x_i, x_j) = (x_i, x_j + c)^d$
Sigmoid	$K(x_i, x_j) = \tan(\sigma(x_i, x_j) + c)$

Cara kerja SVR ditentukan oleh jenis fungsi kernel yang akan dipakai dan pengaturan parameter kernel. Maka pada penelitian ini peneliti akan menggunakan RBF sebagai fungsi kernel SVR. Grid Search Optimization digunakan untuk penentuan nilai parameter. Dalam penggunaannya, Grid Search Optimization biasanya diukur dengan cross validation pada data training. Oleh karena itu disarankan untuk mencoba beberapa variasi pasangan parameter pada hyperplane (Priliani dkk., 2018). Cross validation adalah pengtestingan standar yang dilakukan untuk memprediksi error rate. Data Training dibagi secara random kedalam beberapa bagian dengan perbandingan yang sama kemudian error rate dihitung bagian demi bagian, selanjutnya hitung rata-rata seluruh error rate untuk mendapatkan error rate secara keseluruhan.

Salah satu metode teknik cross validation yang umum digunakan adalah k-fold cross validation. Berikut prosedur k-fold cross validation (Han dkk., 2012):

1. Membagi data menjadi k bagian dengan ukuran yang sama
2. k-1 bagian dijadikan data training dan satu bagian dijadikan data testing
3. proses ini dilakukan sebanyak k pengulangan pada setiap kombinasi data testing dan data training.

Salah satu metode untuk menghitung hasil evaluasi suatu model kita dapat menggunakan perhitungan Root Mean Squared Error (Priliani dkk., 2018).

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

Dimana :

- y_i = Data aktual periode t
- \hat{y}_i = Data prediksi periode t
- n = Jumlah data

3. Metode Penelitian

Uraian Teknik analisis data dalam penelitian ini adalah sebagai berikut

1. Melakukan pengambilan data di situs website Bank Indonesia (BI)
2. Melakukan preprocessing data
3. Melakukan Normalisasi data
4. Membuat model regresi dengan metode *Support Vector Regression*
5. Melakukan prediksi dengan metode *Support Vector Regression*
 - a. Penentuan *Input* dan *Output*. Data *input* dan *output* yang digunakan pada penelitian ini adalah data runtun waktu yaitu data Inflasi berdasarkan Indeks Harga Konsumen setiap bulannya dari Januari 2015 sampai dengan Desember 2020 yang diambil melalui *website* Badan Pusat Statistik.
 - b. Menormalisasikan data.
 - c. Menentukan data *training* dan data *testing*. Dalam penelitian ini jumlah data *training* sebanyak 80 % data, dan 20% data sebagai data *test*.
 - d. Pemodelan dengan *Support Vector Regression* pada data *Training*.
 - e. Penentuan nilai parameter pada data *training*. Penentuan nilai parameter C (*cost*), ϵ (*Epsilon*), dan γ (*Gamma*) dengan menerapkan *Grid Search Optimization*.
 - f. Penentuan model terbaik dengan melakukan evaluasi pada parameter terbaik atau parameter

<https://doi.org/10.35580/variasiunm6>

terkecil.

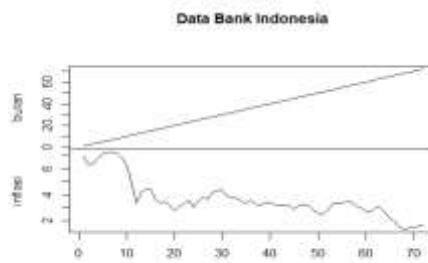
- g. Mengimplementasikan model terbaik pada data *test*.
- h. Melakukan prediksi satu tahun kedepan.

6. Membuat Interpretasi dan kesimpulan berdasarkan hasil pengolahan data:

4. Hasil dan Pembahasan

4.1. Analisis Deskriptif

Analisis deskriptif digunakan untuk melihat Gambaran umum dari data penelitian. Data yang digunakan dalam penelitian ini adalah data Inflasi berdasarkan Indeks Harga Konsumen dari Januari 2015 sampai dengan Desember tahun 2020. Data tersebut diperoleh dari *website* Bank Indonesia(BI). Berikut grafik dan penjelasan dari data tersebut.



Gambar.2 Plot *time series* dari data Inflasi

Dari grafik diatas dapat disimpulkan bahwa selama 6 tahun terakhir dari tahun 2015 – 2020 Inflasi bulanan nasional mengalami penurunan. Dari grafik tersebut dilihat penurunan inflasi hamper setiap bulannya. Penurunan inflasi ini biasa disebut dengan deflasi.

Tabel 2 Statistika Deskriptif data Inflasi Bulanan Nasional

Inflasi	Keterangan (Persen)
Min	1,32
Max	7,26
Median	3,31
Mean	3,66
Std. Deviasi	1,47

Pada Tabel 2 diperoleh nilai minimum dari data Inflasi sebesar 1,32 % dan nilai maksimum sebesar 7,26 %. Standar Deviasi atau jarak antara data terhadap nilai rata – ratanya adalah 1,47 %. Serta Median dari data diperoleh nilai sebesar 3,31 % dan rata – rata dari Inflasi adalah 3,66 %.

Langkah selanjutnya yaitu menormalisasi data menggunakan metode perhitungan normalisasi min – max. Setelah dilakukan normalisasi data, dilanjutkan dengan membagi data menjadi data *train* dan data *test*.

A. Normalisasi Data

Normalisasi data adalah proses pembentukan struktur basis data sehingga sebagian besar *ambiguity* bisa dihilangkan. *Min-Max normalization* merupakan metode normalisasi dengan melakukan

Gambar 4.2 Nilai Inflasi Berdasarkan IHK tahun 2015 - 2020

transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses (Nasution dkk., 2019).

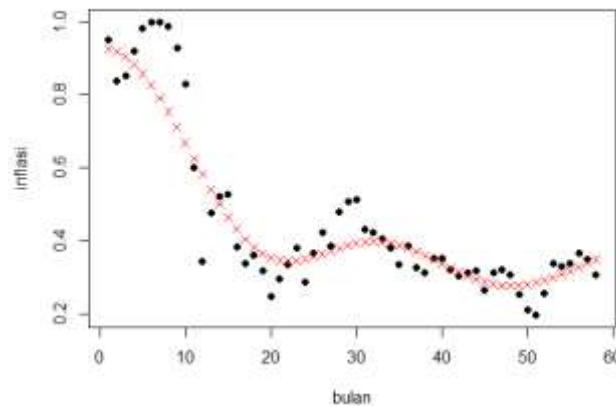
B. Pembagian Data *training* dan Data *testing*

Sebelum data dianalisis menggunakan metode SVR terlebih dahulu data dibagi menjadi data *training* dan data *testing*. Data *train* merupakan data yang digunakan untuk membangun model dengan metode SVR. Sedangkan data *test* adalah data yang digunakan untuk memprediksi dengan metode SVR berdasarkan model yang telah didapatkan sebelumnya.

C. Penerapan Metode Support Vector Regression

1. Membangun Model Support Vector Regression

Untuk membangun model SVR pada *software* R dibutuhkan *package* tambahan yaitu *e1071*. Setelah menginstal *package* *e1071* selanjutnya membuat model SVR pada data *train*. Setelah model didapatkan maka dilakukan prediksi pada data *train*.



Gambar 3 Prediksi SVR tanpa menggunakan kernel

Titik berwarna merah pada Gambar 4.3 merupakan hasil prediksi dengan parameter $\epsilon = 0$, $C = 5$, dan $\gamma = 3$, sedangkan titik hitam merupakan data *train*. Untuk mendapatkan hasil yang maksimal dengan melihat parameter yang terbaik serta *error* paling minim maka diperlukan pengoptimalan model.

2. Pemilihan Model terbaik dengan *Grid Search Optimization*

Pengoptimalan model yang akan digunakan adalah *Grid Search Optimization* dengan tipe kernel *Radial Basis Function*. Pengoptimalan model SVR menggunakan *Grid Search Optimization* membutuhkan parameter sesuai dengan kernelnya. *Grid Search* dijalankan dengan teknik *k - fold cross validation*.

Tabel 3 Parameter hasil *grid search*

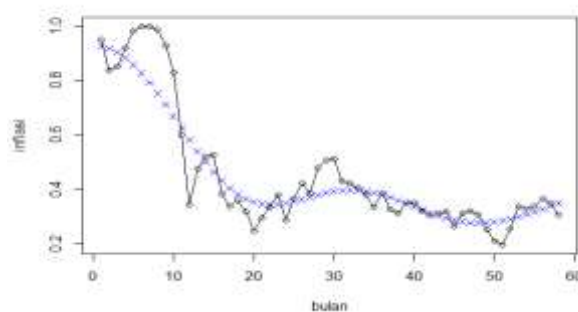
Jumlah <i>k</i>	ϵ Terbaik	C terbaik	γ Terbaik	<i>error</i> terkecil
3	0,1	1	3	0,0046
5	0,1	2	3	0,0057
10	0,1	3	3	0,0055

<https://doi.org/10.35580/variasiunm6>

Didapatkan nilai parameter berbeda – beda untuk nilai k yang berbeda pada Tabel 3. Terlihat bahwa *error* terkecil saat nilai $k = 3$ ketika $\epsilon = 0,1$, $C = 1$, dan $\gamma = 3$ sebesar 0,0046, selanjutnya saat nilai $k = 5$ dengan $\epsilon = 0,1$, $C = 2$, dan $\gamma = 3$ *error* terkecilnya sebesar 0,0057, dan saat nilai $k = 10$ dengan $\epsilon = 0,1$, $C = 3$, dan $\gamma = 3$ *error* terkecilnya sebesar 0,0055. Maka dengan melihat nilai *error* terkecil parameter model yang akan digunakan adalah $\epsilon = 0,1$, $C = 1$, dan $\gamma = 3$ dengan nilai k adalah 3.

Sebelum mendapatkan parameter yang ada pada Tabel 4.5. langkah pertama yaitu memasukkan beberapa nilai parameter pada syntax di R studio. Setelah itu secara *default* output menghasilkan parameter – parameter terbaik dengan *error* terkecil.

Parameter dari model terbaik yang didapatkan sebelumnya digunakan untuk memperediksi data *train*, dan menghasilkan prediksi seperti pada Gambar 4.4 .

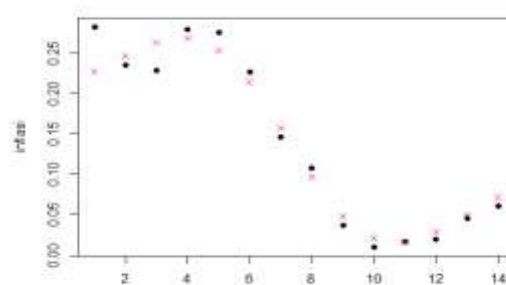


Gambar 4 Prediksi data *Train*

Terlihat pada Gambar 4 hasil prediksi pada data *train* setelah menggunakan parameter terbaik yang telah didapatkan sebelumnya. Hasil prediksi ditunjukkan dengan titik yang berwarna biru, sedangkan data aktual ditunjukkan dengan garis yang berwarna hitam. Dengan melihat Gambar tersebut terlihat bahwa data prediksi mengalami fase naik turun namun tidak drastis seperti pada data aktual.

3. Implementasi pada Data Test

Setelah mendapatkan model terbaik pada langkah sebelumnya. Selanjutnya memasukkan model tersebut beserta parameternya pada data *test*. Hasil prediksi dapat dilihat pada Gambar 4.5



Gambar 5 Plot nilai prediksi pada data *test*

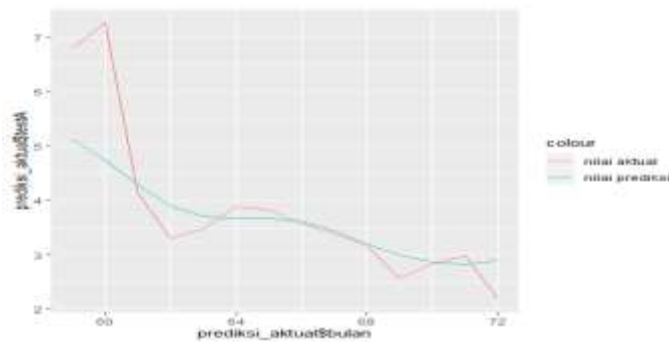
Pada Gambar 5 terlihat hasil prediksi pada data *test* berupa titik berwarna merah sedangkan data *testing* ditunjukkan dengan titik berwarna hitam. Dapat dilihat pada Gambar 4.3 bahwa hasil nilai prediksi mendekati data aktual. Untuk melihat ketepatan hasil prediksi, pada penelitian ini digunakan metode RMSE untuk melihat nilai *sum square error*. Dengan bantuan *software* R hasil perhitungan nilai RMSE

didapatkan nilai 0,0020 yang artinya kemampuan model dapat mengikuti pola data dengan baik. Berikut perbandingan dan nilai aktual pada data *test* yang dapat dilihat pada Tabel 4.7

Tabel 7 Perbandingan Data Aktual dengan Data Prediksi

Bulan	Data aktual	Data prediksi	Bulan	Data aktual	Data prediksi
59	6,79	5,11	66	3,58	3,61
60	7,26	4,74	67	3,40	3,44
61	4,14	4,28	68	3,18	3,20
62	3,31	3,90	69	2,57	2,99
63	3,49	3,70	70	2,83	2,86
64	3,88	3,66	71	2,98	2,83
65	3,82	3,67	72	2,19	2,89

Dari Tabel 7 terlihat bahwa nilai prediksi mendekati nilai aktualnya, sehingga dapat dikatakan bahwa model dengan parameter yang digunakan cocok dengan data yang digunakan. Berikut diagram garis perbandingan data aktual dengan data prediksi.



Gambar 6 Diagram Garis Perbandingan Data Aktual dengan Data Prediksi

Terlihat pada Gambar 6 bahwa data prediksi dengan model terbaik menghasilkan nilai prediksi yang mendekati data aktual.

4. Hasil prediksi Inflasi Bulanan Nasional Indonesia

Hasil prediksi Inflasi di Indonesia dari periode Januari 2021 – Desember 2021 dengan model terbaik dapat dilihat pada Tabel 4.8. Berdasarkan Tabel 4.8 terlihat nilai Inflasi pada bulan Januari sampai dengan Desember 2021 mengalami penurunan setiap bulannya.

6.

Tabel 8 Hasil Prediksi Inflasi di Indonesia

Bulan	Prediksi Inflasi	Bulan	Prediksi Inflasi
Januari 2021	4,28	Juli 2021	3,44
Februari 2021	3,90	Agustus 2021	3,20
Maret 2021	3,70	September 2021	2,99
April 2021	3,66	Oktober 2021	2,86
Mei 2021	3,67	November 2021	2,83
Juni 2021	3,61	Desember 2021	2,89

5. Kesimpulan

Berdasarkan penelitian dan pembahasan yang telah dilakukan untuk memprediksi laju Inflasi di Indonesia dengan metode SVR didapatkan beberapa kesimpulan.

1. Prediksi laju inflasi Indonesia dengan metode Support Vector Regression digunakan metode Grid Search Optimization untuk mendapatkan model terbaik dengan melihat parameter terbaik yaitu $\epsilon = 0,1$, $C = 1$, dan $\gamma = 3$. Model ini akan digunakan untuk memprediksi laju Inflasi satu tahun kedepan.
2. Akurasi hasil prediksi dilakukan dengan melihan nilai Root Mean Square Error. Terlihat nilai RMSE yang dihasilkan pada penelitian ini sebesar 0,0020 yang artinya kemampuan model dapat mengikuti pola data dengan baik.
3. Hasil prediksi Inflasi dari Januari 2021 sampai dengan Desember 2021 dapat dilihat pada Tabel . Terlihat laju Inflasi mengalami fase naik turun namun tidak drastis. Berikut Tabel perbandingan prediksi laju Inflasi Indonesia dengan data aktual Inflasi bulanan Nasional Indonesia.

References

- Ahmar, A. S., Tiro, M. A., Annas, S., dkk. (2018). Modeling Data Containing Outliers using ARIMA Additive Outlier (ARIMA-AO). *Journal of Physics: Conference Series*, 954(1). <https://doi.org/10.1088/1742-6596/954/1/012010>
- Amnda, R. dkk. (FSM U. (2014). Analisis Support Vector Regression (SVR) dalam Memprediksi Kurs Rupiah terhadap Dollar Amerika Serikat. *Jurnal Gaussin*, 03, 849–857.
- Caraka, R. E. (2017). Peramalan Crude Palm Oil (CPO) Menggunakan Support Vector Regression Kernel Radial Basis. 7(1), 43–57.
- Desvina, A., Pani Di, K., Pekanbaru, K., Matematika, J., Sains, F., Sultan, U. I. N., Kasim, S., Hr, J., No, S., & Baru, S. (2015). Penerapan Metode Box-Jenkins dalam Meramalkan Indeks Harga KonsumendDi Kota Pekanbaru. *I(1)*, 39–47.
- Furi, R. P., Si, M., & Saepudin, D. (2015). Prediksi Financial Time Series Menggunakan Independent Component Analysis dan Support Vector Regression Studi Kasus : IHSB dan JII. *ISSN : 2355-9365 e-Proceeding of Engineering :*, 2(2), 1–10.

<https://doi.org/10.35580/variasiunm6>

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining : Concepts and Solution Manual*.

Indonesia, B. (2020). *Data Inflasi*. 2020.

Khatimi, H., & Alkaff, M. (2017). Penerapan support vector regression (svr) untuk peramalan inflasi bulanan nasional. 29–34.

Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. (1997). *Forecasting: Methods and Applications*, Third Edition. null, null.

Maulana, Noval Dini, D. (2019). Implementasi Metode Support Vector Regression (SVR) Dalam

Peramalan Penjualan Roti (Studi Kasus : Harum Bakery). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3), 2986–2995.

Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and System Journal*, 4(1), 78. <https://doi.org/10.24114/cess.v4i1.11458>

Priliani, E. M., Putra, A. T., & Muslim, M. A. (2018). Forecasting Inflation Rate Using Support Vector Regression (SVR) Based Weight Attribute Particle Swarm Optimization (WAPSO). *Scientific Journal of Informatics*, 5(2), 118–127. <https://doi.org/10.15294/sji.v5i2.14613>

Septiningrum, L., Yasin, H., & Sugito. (2015). Prediksi Indeks Harga Saham Gabungan menggunakan Support Vector Regression (SVR) dengan Algoritma Grid Search. 4, 315–321.

Smola, A. J., & Sch, B. (2004). Smola, Schölkopf - 2004 - Statistics and Computing - A tutorial on support vector regression.pdf. *Statistics and Computing*, 14(3), 199–222.

Widiastuti, N. I., Rainarli, E., & Dewi, K. E. (2017). Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen. *Jurnal Infotel*, 9(4), 416. <https://doi.org/10.20895/infotel.v9i4.312>