

IMPLEMENTASI SUPPORT VECTOR REGRESSION PADA PREDIKSI INFLASI INDEKS HARGA KONSUMEN

by Bakhtiyar Hadi Prakoso

Submission date: 21-Apr-2021 12:29PM (UTC+0700)

Submission ID: 1565389794

File name: Bahtiar.pdf (995.19K)

Word count: 4073

Character count: 23003

IMPLEMENTASI SUPPORT VECTOR REGRESSION PADA PREDIKSI INFLASI INDEKS HARGA KONSUMEN

Bakhtiyar Hadi Prakoso

Program Studi Rekam Medik, Jurusan Kesehatan, Politeknik Negeri Jember
Jl. Mastrip, Jember, Jawa Timur
E-mail : bakhtiyar.hp@polije.ac.id

ABSTRAK

Inflasi mencerminkan kenaikan harga barang-barang hal ini sekaligus yang digunakan oleh pemerintah Indonesia khususnya Bank Indonesia dalam menentukan kebijakan moneter. Indikator yang dapat oleh Bank Indonesia dalam mengukur tingkat inflasi adalah Indeks Harga Konsumen. Penelitian ini membahas prediksi inflasi dengan metode SVR. data uji nilai Inflasi yang dikeluarkan oleh Bank Indonesia. Sebagai bahan perbandingan kernel yang digunakan pada metode SVR menggunakan dua kernel yaitu Linear dan Radial Basis Function. Hasil evaluasi tingkat kesalahan menunjukkan bahwa kernel linear menghasilkan nilai yang lebih baik, yaitu dengan tingkat MAPE sebesar 8,70% dan MSE sebesar 0.0037

Kata Kunci : SVM, RBF, Linear, Inflasi

ABSTRACT

Inflation reflects an increase in the prices of these items as well as those used by the Indonesian government, especially Bank Indonesia, in determining monetary policy. An indicator that can be obtained by Bank Indonesia in measuring inflation is the Consumer Price Index. This study discusses inflation prediction using the SVR method. Inflation test data issued by Bank Indonesia. As a comparison material for the kernel used in the SVR method using two kernels, namely Linear and Radial Base Function. The error rate evaluation results show that linear kernels produce better values, with a MAPE rate of 8.70% and MSE of 0.0037

Key word : SVM, RBF, Linear, Inflation

I. PENDAHULUAN

Fungsi utama dari Bank Indonesia adalah menjaga kestabilan nilai rupiah. Untuk mencapai tujuan tersebut, maka Bank Indonesia memerlukan formulasi kebijakan moneter yang efektif untuk mengendalikan inflasi. Inflasi merupakan sebuah kondisi yang mencerminkan kenaikan harga barang-barang. Faktor yang mempengaruhi tingkat inflasi suatu negara biasanya dipengaruhi oleh kondisi dalam negeri. Indikator yang sering digunakan oleh Bank Indonesia dalam mengukur tingkat inflasi adalah Indeks Harga Konsumen [1].

Prediksi merupakan sebuah metode yang dipergunakan untuk menentukan suatu nilai atau kebutuhan di periode selanjutnya. Didalam bank metode ini dapat digunakan untuk membantu manajemen bank dalam mengambil sebuah keputusan terkait kebijakan. Asare menggunakan metode prediksi untuk memprediksi respon pelanggan terkait dengan kampanye promosi. Hasil menunjukkan dapat target yang efektif untuk kampanye promosi [2].

Arutjothi menggunakan KNN untuk memprediksi status pinjaman dengan mengkombinasikan *min-max normalization*, tingkat akurasi dari penelitian tersebut sangat bagus. Lain halnya penelitian yang dilakukan oleh Carvalho, Carvalho menggunakan metode *Gradient Bost Machine* untuk menghitung perkiraan waktu teler bank. Hasilnya mempunyai tingkat akurasi mencapai 97%[3].

Metode-metode yang telah disebutkan sebelumnya merupakan bagian dari metode *machine learning*. Saat ini pemakaian machine learning meluas diberbagai bidang karena keunggulannya. Keunggulan dari *machine learning* dapat dipakai untuk memecahkan permasalahan asosiasi, klusterisasi, klasifikasi, maupun prediksi dalam sebuah himpunan data yang besar.

Saat ini terdapat metode machine learning yang sering dipakai untuk yaitu metode ANN (*Artificial Neural Network*) dan SVM(*Support Vector Machine*) untuk menyelesaikan klasifikasi atau SVR (*Support Vector Regression*) untuk menyelesaikan

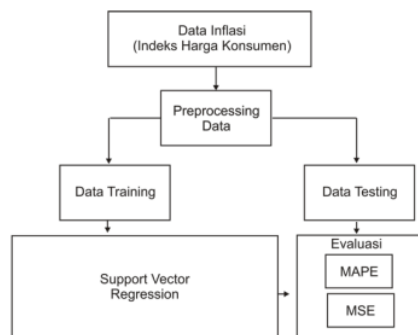
permasalahan prediksi. Kedua metode ini sering dipakai untuk menyelesaikan permasalahan-permasalahan linear maupun non linear dalam kasus prediksi maupun klasifikasi. ANN (*Artificial Neural Network*) dari sisi penggunaan terdapat kelemahan yaitu ketidakmampuannya untuk mengatasi masalah *overfitting* [4]. Lain halnya dengan ANN, SVR (*Support Vector Regression*) muncul sebagai alternatif dari penggunaan ANN. SVR (*Support Vector Regression*) turunan metode (*Support Vector Machine*) yang telah dimodifikasi [5] untuk keperluan prediksi [6]. SVM mempunyai target value bernilai *Boolean*. Sedangkan SVR mempunyai target value bernilai kontinyu. Salah satu kelebihan metode SVR/SVM dari pada ANN adalah dapat menyelesaikan permasalahan *overfitting* dan tingkat akurasi yang lebih tinggi [7].

Kelebihan SVR dibandingkan dengan ANN membuat metode SVR banyak dipakai. Beberapa penelitian yang telah menggunakan SVR diantaranya prediksi konsumsi elektrik [8], prediksi beban mesin [9], prediksi durasi dari proyek perangkat lunak [10]. Dari ketiga penelitian tersebut, semuanya memberikan hasil prediksi yang baik.

Penelitian ini membahas implementasi metode SVR untuk memprediksi data inflasi berdasarkan Indeks Harga Konsumen. Data yang diambil berasal dari Bank Indonesia yang didapatkan melalui situs Bank Indonesia periode maret 2003 sampai dengan Juli tahun 2017.

II. METODOLOGI

Kerangka konsep penelitian didalam penelitian ini adalah sebagai berikut



Gambar 1. Kerangka Konsep

Data yang dipakai dalam penelitian ini adalah

data inflasi Indeks Harga konsumen dimana data diambil dari situs resmi Bank Indonesia pada periode Maret 2003 sampai dengan Juli 2017. yang didapatkan dari Bank Indonesia terlebih dahulu melalui tahapan *preprocessing* data agar dapat digunakan untuk tahapan selanjutnya. Dari proses *preprocessing* data dibagi menjadi dua jenis yaitu data uji dan data latih, kemudian data digunakan untuk proses uji.

Metode SVR akan dilakukan uji terhadap dua kernel yang berbeda yaitu kernel *Radial Basis Function* dan kernel *Linear*. Kedua kernel ini diambil karena dibebberapa penelitian dapat menghasilkan nilai prediksi yang baik. Sedangkan untuk proses evaluasi digunakan *MAPE (Mean Absolute Presentage Error)* dan *MSE (Mean Square Error)*

1. Inflasi

Inflasi merupakan salah satu indeks acuan yang menggambarkan kenaikan harga barang yang terjadi secara terus menerus. Istilah inflasi bukan berarti naiknya harga satu atau dua barang saja, namun kenaikan harga barang-barang secara meluas. Indikator inflasi pada umumnya didasarkan pada [11]

1. Indeks Harga Perdagangan Besar (IHPB). Merupakan harga transaksi yang terjadi antara penjual/pedagang besar pertama dengan pembeli/pedagang besar berikutnya dalam jumlah besar pada pasar pertama atas suatu komoditas.
2. Deflator Produk Domestik Bruto (PDB) merupakan hasil pengukuran level harga barang akhir dan jasa yang diproduksi di dalam negeri.

Indikator yang sering digunakan oleh Bank Indonesia untuk mengukur tingkat inflasi adalah Indeks Harga Konsumen (IHK). Nilai IHK dari tahun ke tahun akan mengalami pergerakan. Pergerakan ini dapat dimonitor dari pergerakan harga paket barang dan jasa yang dikonsumsi oleh masyarakat. Indeks Harga Konsumen (IHK) ditentukan berdasarkan survei yang diadakan oleh Badan Pusat Statistik. Atribut survei yang digunakan adalah Survei Biaya Hidup. Survei ini dimonitor di beberapa kota dalam interfal waktu satu bulan dan dilaukan di pasar tradisional atau pasar modern terhadap beberapa jenis barang dan jasa.

2. Time Series Analysis

Time series merupakan sekumpulan data

yang diobservasi selama kurun waktu tertentu yang terurut secara kronologis. Biasanya data-data *time series* akan dianalisis untuk mengetahui tingkah laku data, menghitung estimasi dan prediksi serta mengamati beberapa faktor-faktor yang mempengaruhinya. Dalam analisis *time series* topik yang sering dibahas adalah tentang prediksi/peramalan [11]. Data *time series* yang diobservasi dapat diproses sehingga didapatkan prakiraan untuk waktu selanjutnya. Jika ditinjau dari jenis datanya analisis *time series* dibagi menjadi dua yaitu

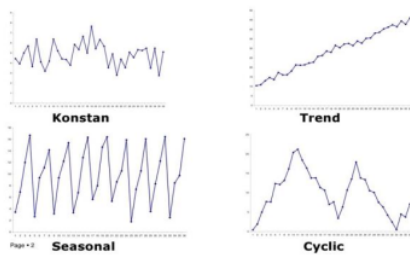
a) *Univariate time series*

Data *univariate time series* merupakan data yang hanya mempunyai satu buah variabel / atribut yang mempengaruhi nilai

b) *Multivariate time series*

Data *multivariate time series* merupakan data yang mempunyai beberapa buah variabel yang mempengaruhi nilai

Salah satu aspek penting dalam memilih metode prediksi adalah melihat pola dari data yang diobservasi. Didalam prediksi terdapat empat macam pola yang diperkenalkan yaitu sebagai [10] :



Gambar 2 Pola Data Time Series

- a) Pola Horizontal / Konstan
Pola horizontal merupakan pola ketika nilai dari data cenderung fluktuatif disekitar nilai mean.
- b) Pola Musiman / Seasonal
Pola dimana data *time series* dipengaruhi oleh data musiman. Sebagai contoh di Eropa Baju tebal akan laku keras pada bulan Nopember dan Desember.
- c) Pola Siklis
Pola siklis merupakan pola dimana *time series* mengalami kenaikan dan penurunan yang tidak mempunyai pola.
- d) Pola Trend
Pola *time series* dimana data mengalami

kenaikan maupun penurunan dalam kurun waktu yang lama

3. *Support Vector Regression*

Support Vector Regression merupakan hasil dari modifikasi dari metode *Support Vector Maching* yang dipergunakan untuk menyelesaikan masalah prediksi [12]. Dalam konsep regresi fungsi *linear* secara umum dapat dituliskan sebagai berikut

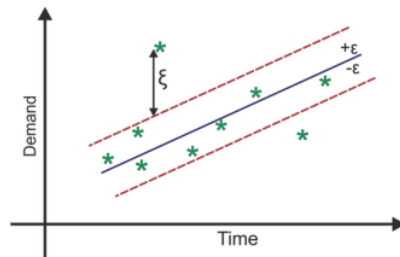
$$f(x) = \langle w, x \rangle + b \quad w \in \mathcal{X}, b \in \mathbb{R} \quad (1)$$

Di mana $\langle ., . \rangle$ hasil dari *dot product* di \mathcal{X} . Fungsi 2.1 merupakan fungsi *flatness* untuk mencari nilai w yang paling kecil. Nilai tersebut dapat dicari dengan meminimalkan bentuk $\|w\|^2$. Permasalahan tersebut dapat diselesaikan dengan *convex optimization* yang dijelaskan pada rumus 2.

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \|w\|^2 \quad (2) \\ \text{Subject to :} \quad & y_i - \langle w, x \rangle - b \leq \varepsilon \\ & \langle w, x \rangle + b - y_i \leq \varepsilon \end{aligned}$$

Terdapat kondisi dimana nilai kesalahan melebihi batas ambang ε . Pada kondisi ini dibutuhkan *soft margin* atau bisa disebut dengan *variable slack* ξ_i, ξ_i^* . Sehingga persamaannya berubah seperti pada rumus 3.

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=0}^{\ell} (\xi_i + \xi_i^*) \\ \text{Subject to :} \quad & y_i - \langle w, x \rangle - b \leq \varepsilon + \xi_i \\ & \langle w, x \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \quad (3) \end{aligned}$$



Gambar 3 Konsep *Support Vector Regression (SVR)*

Sumber : Levis dkk[7]

Konstanta $C > 0$ menentukan seberapa besar tingkat deviasi kesalahan dari batas ε yang dapat ditolerasi.

Rumus diatas merupakan permasalahan

Convex Linear Programing NLP Optimization Problem yang berfungsi untuk meminimalkan fungsi kuadrat untuk diubah menjadi sebuah batasan. Batasan tersebut dapat diselesaikan dengan persamaan *Lagrange Multiplier*. Proses penurunan rumus sangat panjang dan rumit. Setelah melalui tahapan-tahapan matematis maka didapatkan persamaan baru dengan rumus rumus 4.

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \cdot \langle x_i, x \rangle + b \quad (4)$$

Di mana x_i merupakan *support vector* dan x merupakan *test vector*. Fungsi diatas dapat digunakan untuk menyelesaikan permasalahan *linear*. Sedangkan untuk permasalahan *non linear* nilai x_i dan x terlebih dahulu ditransformasikan ke dalam *space feature* yang berdimensi tinggi dengan cara memetakan vektor x_i dan x ke dalam fungsi kernel Sehingga persamaan akhirnya menjadi :

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b. \quad (5)$$

Fungsi $K(x_t, x_t)$ merupakan *Kernel*. Tabel dibawah ini merupakan kernel yang dipakai dalam perhitungan SVR (Levis & Papageorgiou, 2005)

Tabel 1. Kernel Untuk SVR

No	Nama Kernel	Formula
1	Linear	$K(x_i, x) = x_i \cdot x$
2	Radial Basis Function	$K(x_i, x) = \exp\left(\frac{\ x_i - x\ ^2}{-2\sigma^2}\right) \gamma = \frac{1}{2\sigma^2}$

4. MAPE dan MSE

MAPE (Mean Presentage Error) merupakan teknik pengukuran tingkat kesalahan yang paling popular dalam prediksi. MAPE dihitung dari rata-rata abolute dari presentage error. Misalkan A_t dan F_t berturut turut Nilai actual dan nilai prediksi dan . Sedankan N didefinisan sebagai jumlah periode yang diobservasi Maka MAPE didefinisiakan sebagai [13]

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{A_t - F_t}{A_t} \right| \quad (6)$$

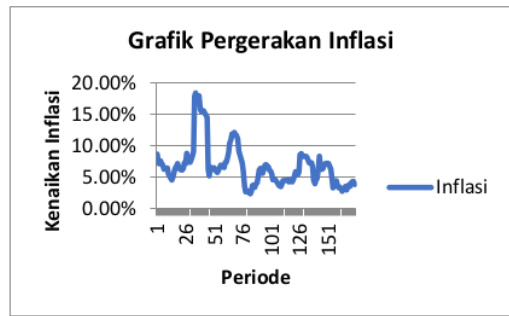
Sedangkan MSE (*Mean Square Error*) digunakan untuk mengukur tingkat akurasi berdasarkan quadratic loss [14]

$$MSE = \frac{1}{N} \sum_{t=1}^N (A_t - F_t)^2 \quad (7)$$

Nilai MSE sangat sensitive terhadap nilai kesalahan atau outlier. Jika terdapat salah satu outlier saja nilai MSE akan menjadi besar.

III. HASIL DAN PEMBAHASAN

Data yang dipakai penelitian adalah data inflasi dari Bank Indonesia pada bulan Maret 2003 sampai dengan Juli 2017. Jenis data inflasi yang digunakan adalah Indeks Harga Konsumen karena jenis data inflasi ini paling sering digunakan oleh Bank Indonesia dalam mengukur tingkat inflasi. Data inflasi dalam bentuk data *time series* dengan interfal waktu per satu bulan. Grafik pergerakan inflasi dapat ditunjukkan pada gambar 4



Gambar 4 Data inflasi periode Maret 2003 sampai dengan Juli 2017

Data yang didapat dalam bentuk persentase. Deskripsi data dapat dijelaskan sebagai berikut nilai terendah data adalah 2,410%, nilai kuartal pertama sebesar 4,515%, nilai median sebesar 6,290% nilai kuartal ketiga sebesar 6,290% dan nilai tertinggi sebesar 1,838%. Jika dilihat pada bentuk grafiknya pola *time series* cenderung berpola siklis dimana data tidak mempunyai pola.

Data yang dipergunakan melewati proses *preprocessing data*. Proses ini bertujuan agar model yang dihasilkan dari perhitungan Support Vector Regression dapat menghasilkan tingkat akurasi yang baik. *Preprocessing Data* dimulai dari merubah bentuk data dari persentase kedalam bentuk decimal, tujuannya adalah untuk memudahkan perhitungan.

Penelitian ini menggunakan jenis Proses selanjutnya adalah penyusunan data, kedalam format *Support Vector Regression*. Format penyusunan terbagi menjadi dua yaitu target value dan atribut. Nilai value merupakan hasil nilai prediksi sedangkan nilai atribut merupakan nilai-nilai yang berpengaruh terhadap nilai prediksi. Didalam penelitian ini adalah data lag dari atribut yang berpengaruh adalah 3 langkah mundur kebelakang. Dimisalkan nilai *target value* adalah x_t maka nilai dari masing-masing atribut adalah x_{t-1} , x_{t-2} , dan x_{t-3} Hasil pemodelan ditunjukkan pada rumus 10.

$$X = \begin{bmatrix} x_{t-3} & x_{t-2} & x_{t-1} \\ x_{t-2} & x_{t-1} & x_t \\ \vdots & \vdots & \vdots \\ x_{t-3+n} & x_{t-2+n} & x_{t-1+n} \end{bmatrix}, Y = \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_n \end{bmatrix} \quad (10)$$

Adapun langkah-langkah preprocessing data adalah sebagai berikut

- 1) Merubah data : data yang didapat merupakan nilai dalam bentuk persentase, data ini dirubah terlebih dahulu ke dalam bentuk pecahan untuk memberikan kemudahan dalam menghitung nilai prediksi
- 2) Penyusunan data kedalam format data input *Support Vector Regression*. Format penyusunan terbagi menjadi dua yaitu target value dan atribut. Asumsi yang dipakai didalam penelitian ini adalah data lag dari atribut yang berpengaruh adalah 3 langkah mundur kebelakang. Dimisalkan nilai *target value* adalah x_t maka nilai dari masing-masing atribut adalah x_{t-1} , x_{t-2} , dan x_{t-3} Hasil pemodelan ditunjukkan pada rumus 10.

$$X = \begin{bmatrix} x_{t-3} & x_{t-2} & x_{t-1} \\ x_{t-2} & x_{t-1} & x_t \\ \vdots & \vdots & \vdots \\ x_{t-3+n} & x_{t-2+n} & x_{t-1+n} \end{bmatrix}, Y = \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_n \end{bmatrix} \quad (10)$$

- 3) Dari pemodelan yang dilakukan Data inflasi secara keseluruhan akan dibagi menjadi dua: data latih dan data uji. Data latih diambil dari periode waktu . Sedangkan data uji diambil dari periode waktu Juli 2016 sampai dengan Juli 2017

Pengujian dengan kernel *Radial Basis Function* dipengaruhi oleh dua parameter utama yaitu γ dan C sedangkan untuk nilai δ diasumsikan dengan nilai 0.1 yang merupakan nilai default yang ada didalam R. Proses

penentuan parameter γ dan C cukup sulit karena terdiri dari dua parameter Nilai γ dan C yang paling optimal ditentukan dengan algoritma grid search, algoritma ini merupakan fitur yang ada didalam R. Berikut ini adalah code program yang digunakan untuk mencari parameter terbaik γ dan C.

```
#pencarian parameter terbaik
#=====
> best.model = tune.svm(y~., data=datatrain, cost =
10^(-1:3), gamma = 10^(-1:3), kernel="radial")
> best.model
Parameter tuning of 'svm':
- sampling method: 10-fold cross validation
- best parameters:
  gamma cost
  0.1 10
- best performance: 0.000141997
```

Nilai cost dan gamma diambil dalam rentang 10^{-1} sampai dengan 10^3 . Dari kode program tersebut didapatkan nilai γ dan C terbaik berturut turut adalah 0.1 dan 10. Berikut ini adalah code program untuk proses menghitung nilai prediksi.

```
#perhitungan prediksi
#=====
> model.rbf = svm(y~., data=datatrain, type='eps-
regression', kernel='radial', gamma=0.1, epsilon=0.1,
cost=10, scale = TRUE)
> prediction.rbf = predict(model.rbf, datatest[-1])
```

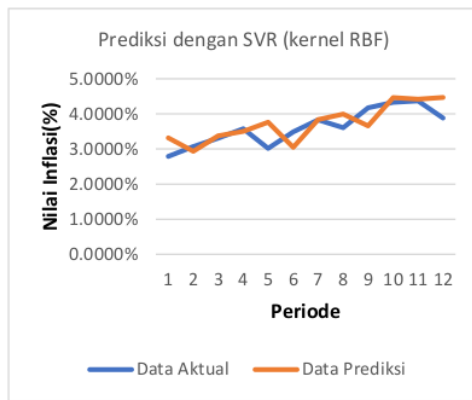
Dari kode program diatas, didapatkan hasil prediksi dengan menggunakan kernel *Radial Basis Function* dengan parameter $\gamma=0.1$ dan C = 10 adalah sebagai berikut

Tabel 2 Hasil prediksi kernel RBF

No	Data Aktual	Data Prediksi
1	2.7900%	3.3130%
2	3.0700%	2.9355%
3	3.3100%	3.3772%
4	3.5800%	3.4997%
5	3.0200%	3.7649%
6	3.4900%	3.0465%
7	3.8300%	3.8372%
8	3.6100%	3.9917%
9	4.1700%	3.6618%

10	4.3300%	4.4668%
11	4.3700%	4.4175%
12	3.8800%	4.4682%
MAPE	8.9276%	
MSE	1,52E-05	

Hasil perhitungan berdasarkan algoritma grid search menunjukkan bahwa tingkat kesalahan berdasarkan nilai MAPE sebesar 9.2767% sedangkan nilai MSE sebesar 1,52E-05. Jika digambarkan kedalam bentuk grafik maka akan tampil Gambar 4.



Gambar 5 Hasil prediksi dengan kernel RBF

Gambar 5 merupakan visualisasi hasil prediksi perhitungan dengan menggunakan kernel Radial Basis Function. Garis biru merepresentasikan data aktual, sedangkan garis oranye merepresentasikan data prediksi. Dari gambar tersebut hasil prediksi cukup baik mengikuti tren data.

Uji data dengan SVR (*Support Vector Regression*) kernel *linear* dilakukan 3 kali. Fokus pengujiannya dengan mencari nilai cost yang terbaik, karena nilai ini berpengaruh besar terhadap nilai prediksi sedangkan untuk nilai epsilon diasumsikan sebesar 0.1 nilai ini merupakan nilai default dari R. Parameter Cost ditentukan secara manual yang akan diuji adalah Cost (C) sebesar, 1,10, dan, 100. Berikut ini code program di R yang digunakan untuk memodelkan dan memprediksi nilai inflasi

```
#model
mymodel1 = svm(y~., data=datatrain, type='eps-
regression', kernel='linear', epsilon=0.1, cost=1,
```

```
scale = TRUE);
mymodel2 = svm(y~., data=datatrain, type='eps-
regression', kernel='linear', epsilon=0.1, cost=10,
scale = TRUE);
mymodel3 = svm(y~., data=datatrain, type='eps-
regression', kernel='linear', epsilon=0.1,
cost=100, scale = TRUE);
```

```
#prediksi
myprediction1 = predict(mymodel1, datatest[-
1]);
myprediction2 = predict(mymodel2, datatest[-
1]);
myprediction3 = predict(mymodel3, datatest[-
1]);
```

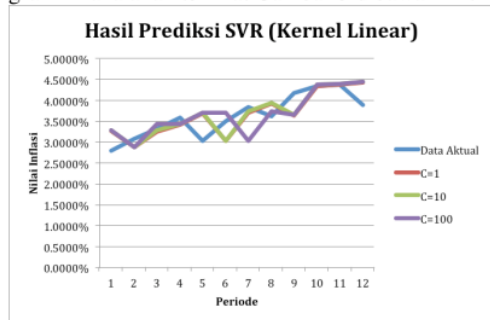
Sebelum dihitung, data dimodelkan terlebih dahulu dengan menentukan tipe dari kernel yang akan digunakan, dan memasukkan parameter-parameter pendukung. Parameter yang dimasukkan adalah nilai dari epsilon dan cost. Pemodelan ini berfungsi sekaligus sebagai proses training data sehingga data yang dimasukkan adalah data latih. Lainnya halnya pada langkah pemodelan. Pada tahap perhitungan prediksi data yang digunakan adalah data tes. Hasil uji coba dapat dilihat pada tabel 3

Tabel 3 Data Hasil Uji Metode SVR dengan Kernel Linear

No	Data Aktual	Data Prediksi		
		C=1	C=10	C=100
1	2.7900%	3.2757%	3.2773%	3.2774%
2	3.0700%	2.8754%	2.8749%	2.8750%
3	3.3100%	3.2608%	3.2820%	3.4405%
4	3.5800%	3.4259%	3.4404%	3.4405%
5	3.0200%	3.6951%	3.7104%	3.7106%
6	3.4900%	3.0335%	3.0248%	3.7106%
7	3.8300%	3.7065%	3.7338%	3.0249%
8	3.6100%	3.9268%	3.9421%	3.7340%
9	4.1700%	3.6464%	3.6466%	3.6467%
10	4.3300%	4.3477%	4.3743%	4.3745%
11	4.3700%	4.3777%	4.3865%	4.3867%
12	3.8800%	4.4348%	4.4430%	4.4432%
	MAPE	8.7010%	8.7445%	9.4816%
	MSE	1.37E-05	1.44E-05	1.76E-05

Tabel 3 menunjukkan bahwa MAPE terkecil terletak pada nilai Cost = 1 dengan nilai 8.7010% dan nilai MAPE terbesar sebesar 9.4816% pada cost = 100. Hasil tersebut menunjukkan bahwa parameter Cost =1 merupakan parameter yang paling maksimal dibandingkan dengan Cost =10, dan Cost = 100 Sedangkan MSE terkecil pada C=1 dengan nilai

1.37E-05 dan MSE terbesar pada C=100 dengan nilai 0.0042. Jika di gambarkan dalam bentuk grafik maka akan terlihat Gambar 5 dibawah ini.



Gambar 5 Grafik perbandingan uji kernel linear

Gambar 5 merupakan hasil visualisasi data hasil uji dalam bentuk grafik. Data aktual direpresentasikan dengan garis biru, sedangkan garis oranye, abu-abu, dan kuning merepresentasikan data hasil uji dengan parameter C = 1, C=10, dan C=100. Jika dilihat dari visual nilai prediksi C=1 dan C=10 nilainya hampir sama di dua belas periode terakhir.

Dengan begitu nilai prediksi terbaik pada metode SVR kernel linear adalah dengan menggunakan konstanta C = 1, sedangkan pola sebaran nilai kesalahan dapat dilihat dengan nilai MSE yang paling kecil yaitu pada C = 1.

Uji prediksi dengan menggunakan dua parameter penilai yaitu MAPE dan MSE, Hasil pengujiannya dapat dirihat pada tabel 4 dibawah ini

Tabel 4 Perbandingan MAPE dan MSE

PARAMETER	SVR (RBF)	SVR (Linear)		
		C=1	C=10	C=100
MAPE	8.93%	8.70%	8.74%	9.48%
MSE	1.52E-05	1.37E-05	1.44E-05	1.76E-05

Tabel 4 mendeskripsikan hasil uji yang telah dilakukan pada data inflasi. Uji dilakukan dengan dengan metode Support Vector Regression. Uji dengan Support Vector Regression dilakukan dengan dua kernel yang berbeda yaitu RBF (Radial Basis Function) dengan kernel linear. Pada pengujian dengan menggunakan metode SVR nilai MAPE terkecil didapatkan dari kernel linear dengan parameter C (Cost) = 1 dan nilai MAPE sebesar 8,6%

sedangkan nilai terbesar pada penggunaan kernel RBF dengan nilai MAPE 8.93%.

IV. SIMPULAN DAN SARAN

Berdasarkan hasil uji yang telah dilakukan dengan dua kernel yang berbeda, maka didapatkan hasil prediksi data inflasi indeks harga konsumen yang mempunyai tingkat kesalahan terendah adalah dengan menggunakan kernel liner dengan nilai C atau cost sebesar 1, dengan nilai MAPE dan MSE berturut turut sebesar 8,70% dan 1.37E-05. Proses penentuan parameter merupakan bagian tersulit dalam tahapan perhitungan metode Support Vector Regression karena terdapat beberapa parameter yang dikombinasikan. Pada penelitian ini penulis menggunakan dua cara yaitu manual otomatis. Pada kernel linear (mencari nilai Cost) nilai cost diasumsikan dengan C=1,10, dan 100. Pada kernel radial basis fuction penentuan parameter menggunakan grid search pada kernel RBF (mencari nilai Cost dan Gamma). Perlu adanya kajian ulang terkait penentuan parameter parameter yang tepat sehingga hasil prediksi metode SVR data ditekan lagi nilai kesalahannya. Selain itu dapat dikaji lagi penggunaan kernel yang lain, selain kernel linear adan Radial Basis Function.

V. UCAPAN TERIMAKASIH

Ucapan terimakasih dihaturkan kepada keluarga yang telah mendukung dan kampus Politeknik Negeri Jember yang telah memberi kesempatan waktu bagi penulis untuk menyelesaikan penelitian ini. Tidak lupa kepada Bank Indonesia yang telah mempublikasi data inflasi di halaman web, sehingga peneliti dapat mendapatkan data dengan mudah.

REFERENSI

- [1] N. Joko Prastowo, T. Yanuarti, and Y. Depari, "PENGARUH DISTRIBUSI DALAM PEMBENTUKAN HARGA KOMODITAS DAN IMPLIKASINYA TERHADAP INFLASI," *Bank Indones.*, 2008.
- [2] J. Asare-Frempong and M. Jayabalan, "Predicting customer response to bank direct telemarketing campaign," *2017 Int. Conf. Eng. Technol. Technopreneurship, ICE2T 2017*, vol. 2017-January, pp. 1-4, 2017.
- [3] R. S. Carvalho, R. N. Carvalho, G. N.

- Ramos, and R. N. Mour^o, "Predicting waiting time overflow on bank teller queues," *Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017*, vol. 2018-January, pp. 842–847, 2018.
- [4] G. P. Zhang, E. B. Patuwo, and H. Michael Y., "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35–62, 1998.
- [5] V. N. Vapnik, "An overview of statistical learning theory.," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, 1999.
- [6] C. Lin and R. C. Weng, "Simple Probabilistic Predictions for Support Vector Regression," *Statistics (Ber.)*, pp. 1–16, 2004.
- [7] A. Levis and L. Papageorgiou, "Customer demand forecasting via support vector regression analysis," *Cemical Eng. Reaserch Des.*, no. August, 2005.
- [8] F. Magoulès, M. Piliouguine, and D. Elizondo, "Support Vector Regression for Electricity Consumption Prediction in a Building in Japan," *Proc. - 19th IEEE Int. Conf. Comput. Sci. Eng. 14th IEEE Int. Conf. Embed. Ubiquitous Comput. 15th Int. Symp. Distrib. Comput. Appl. to Business, Engi.*, pp. 189–196, 2017.
- [9] L. W. Chong, D. Rengasamy, Y. W. Wong, and R. K. Rajkumar, "Load prediction using support vector regression," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2017-Decem, pp. 1069–1074, 2017.
- [10] C. Lopez-Martin, S. Banitaan, A. Garcia-Florianio, and C. Yanez-Marquez, "Support vector regression for predicting the enhancement duration of software projects," *Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017*, vol. 2018-Janua, pp. 562–567, 2018.
- [11] R. Hrasko, G. C. Pacheco, and R. A. Krohling, "Time Series Prediction using Restricted Boltzmann Machines and Backpropagation," *Inf. Technol. Quant. Manag.*, vol. 55, no. Itqm, pp. 990–999, 2015.
- [12] A. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, pp. 199–222, 2004.
- [13] A. Botchkarev, "Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology," pp. 1–37, 2018.
- [14] P. A. Thompson, "An MSE statistic for comparing forecast accuracy across series," *Int. J. Forecast.*, vol. 6, no. 2, pp. 219–227, 1990.

IMPLEMENTASI SUPPORT VECTOR REGRESSION PADA PREDIKSI INFLASI INDEKS HARGA KONSUMEN

ORIGINALITY REPORT

9%

SIMILARITY INDEX

11%

INTERNET SOURCES

11%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Universitas Bina Darma

Student Paper

6%

2

Rizki Rino Pratama. "Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia", MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, 2020

Publication

2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On