

TUGAS

Prediksi Harga Saham JKSE Menggunakan Support Vector Regression

Diajukan untuk memenuhi tugas mata kuliah Analisis Data Deret Waktu II

Dosen Pengampu : Rezzy Eko Caraka, Ph. D



Disusun oleh:

Haana Lahanda Putri Samsi	140610210010
Muhamad Raditya Putra Marindra	140610210017
Muhammad Nauval Muzaki	140610210022
Juzma Fawwaza Rosadi	140610210055
Muhammad Khatami Sahl	140610210079

Program Studi Statistika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Padjadjaran

2023

1. Pendahuluan

Regresi merupakan salah satu metode pada analisis data dalam statistika yang digunakan untuk melakukan peramalan, maupun mengkaji hubungan antar variabel. Terdapat dua jenis analisis regresi atau biasa disebut regresi linear, yaitu regresi linear sederhana dan regresi linear berganda.

Terdapat contoh lain dari jenis metode regresi yaitu *Support Vector Regression* (SVR) yang merupakan bentuk regresi dari metode *Support Vector Machine* (SVM). Jika SVM untuk masalah klasifikasi dengan variabel dependen kategorik, SVR digunakan pada kasus regresi yang variabel dependennya berupa bilangan riil atau kontinu.

Data yang digunakan untuk penelitian ini bersumber dari website Yahoo Finance yang merupakan data *closing price* dari indeks Saham *Jakarta Stock Exchange* (JKSE) sebanyak 276 data. Pemodelan dan prediksi dilakukan dengan melibatkan seluruh data serta membagi data ke bagian training dan testing. Analisis data dilakukan dengan metode *Support Vector Regression* (SVM) menggunakan *Software R*.

2. Metodologi Penelitian

Support Vector Regression (SVR)

Support Vector Regression (SVR) adalah penyesuaian dari teori *machine learning* yang sebelumnya digunakan untuk memecahkan masalah klasifikasi, yaitu *Support Vector Machine* (SVM). Dalam kata lain, SVM untuk kasus regresi disebut sebagai SVR. Tujuan dari SVR adalah untuk menemukan sebuah fungsi $f(x)$ sebagai suatu *hyperplane* berupa fungsi regresi yang sesuai dengan semua input data dengans sebuah *error* ε dan membuat ε setipis mungkin. Fungsi regresi dari metode SVR dituliskan sebagai berikut.

$$f(x) = \sum_{i=1}^{n^*} w_i \cdot \varphi_i(x) + b$$

$$w = \sum_{i=1}^{n^*} (\alpha_i - \alpha_i^*) x_i$$

Data *training* yang memiliki nilai $\alpha_i > 0$ adalah data yang terpilih menjadi *support vector*, sedangkan sisanya memiliki nilai 0. Dengan demikian fungsi keputusan yang dihasilkan hanya dipengaruhi oleh *support vector*.

Fungsi Kernel

Pada dunia nyata kasus-kasus yang ditemukan adalah kasus yang tidak linear sehingga untuk mengatasi masalah ketidaklinearan dapat diterapkan fungsi kernel. Berikut ini merupakan beberapa fungsi kernel menurut Hsu et al.

1. Kernel Linear

$$K(x_i, x_j) = x_i^T x_j$$

2. Kernel Polinomial

$$K(x_i, x_j) = (\mu x_i^T x_j + r)^d, \mu > 0$$

3. Kernel Radial Basis Function (RBF)

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

4. Kernel Sigmoid

$$K(x_i, x_j) = \tanh(\mu x_i^T x_j + r)$$

Mean Percentage Error (MAPE)

MAPE merupakan teknik pengukuran tingkat kesalahan yang paling populer dalam prediksi. MAPE dihitung dari rata-rata absolut dari persentase error. Misalkan A_t dan F_t berturut-turut nilai aktual dan nilai prediksi. Sedangkan N didefinisikan sebagai jumlah periode yang diobservasi maka MAPE didefinisikan sebagai

$$MAPE = \frac{1}{n} \sum_{t=1}^N \left| \frac{A_t - F_t}{A_t} \right|$$

R-Square

R-Square merupakan suatu nilai yang memperlihatkan seberapa besar variabel independen memengaruhi variabel dependen. *R-Square* merupakan angka yang berkisar antara 0 sampai 1 yang mengindikasikan besarnya kombinasi variabel independen secara bersama-sama memengaruhi nilai variabel dependen.

3. Hasil dan Pembahasan

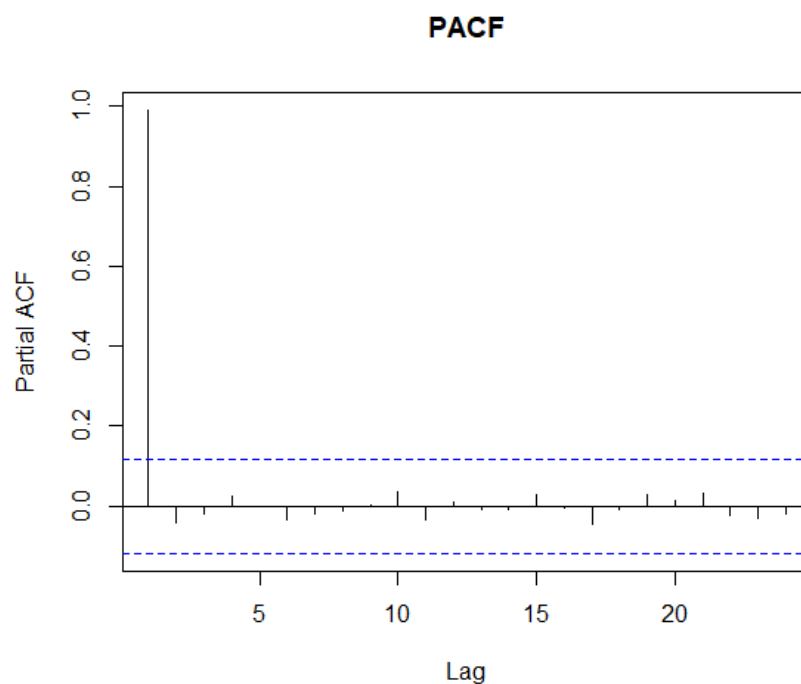
3.1. Statistika Deskriptif

Tabel 1. Statistika Deskriptif

	Mean	StdDev	Minimum	Maximum
JKSE Close Prices	3621,5	2225,981	358,2	7228,9

Berdasarkan Tabel 1 dapat dilihat bahwa harga minimum saham JKSE adalah 358,2. Kemudian harga maksimum saham JKSE adalah 7228,9. Harga saham JKSE memiliki keragaman yang tinggi dengan nilai varian (Standar Deviasi) sebesar 2225,981. Harga saham JKSE memiliki rata-rata harga saham sebesar 3621,5. Identifikasi lag sebagai komponen input didasarkan pada lag-lag yang memiliki nilai PACF terbesar. Hal ini dikarenakan karakteristik persamaan model tersebut memiliki kesamaan dengan model AR.

3.2. *Partial Autocorrelation Function*



Gambar 1. *Partial Autocorrelation Function*

Berdasarkan gambar 1. Dapat dilihat bahwa lag keluar dari garis stasioner pada lag ke-1 oleh karena itu dalam jurnal kami menggunakan lag 1 atau X_{t-1} .

3.3. **Parameter Terbaik**

Selanjutnya akan dilakukan estimasi parameter dan mendapatkan parameter terbaik dari setiap kernel. Dengan menggunakan *Cost* 0,1, 1, 10, 100 dan *Epsilon*(ϵ) 0,1, 0,01, 0,001, 0,001.

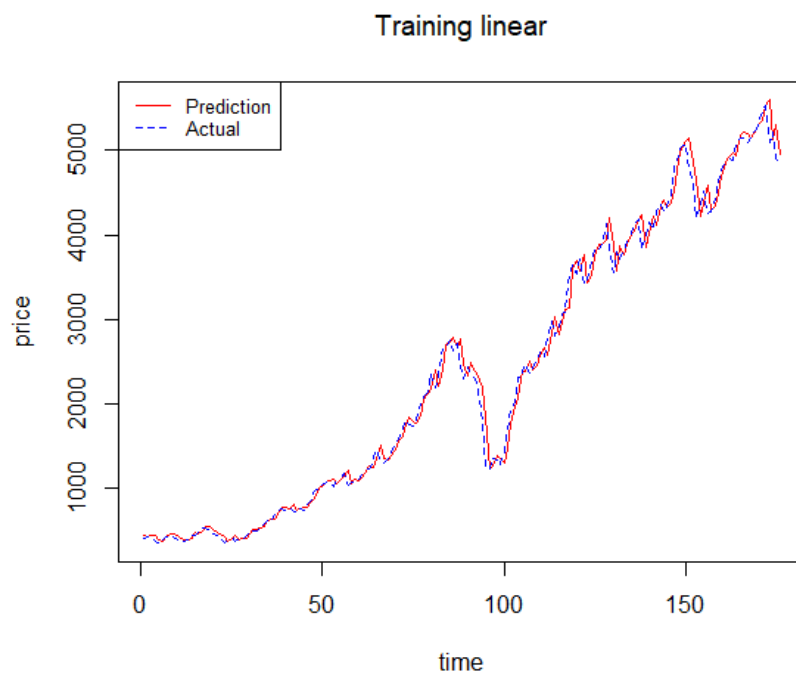
Tabel 2. Estimasi Parameter

Kernel	<i>Best Cost</i>	<i>Best Epsilon</i> (ϵ)
Linear	100	0,01
Polynomial	1	0,1
Sigmoid	0,1	0,1
Radial	1	0,01

Berdasarkan tabel 2. Didapatkan *Cost* dan *Epsilon*(ϵ) terbaik dari setiap kernel yang nantinya akan dibandingkan hasil dari setiap kernel dengan akurasi *R-Square* dan MAPE(*Mean Absolute Percentage Error*).

3.4. Plot Prediction and Actual

3.4.1. Linear Kernel



Gambar 2. Plot Prediction and Actual Training with Linear Kernel

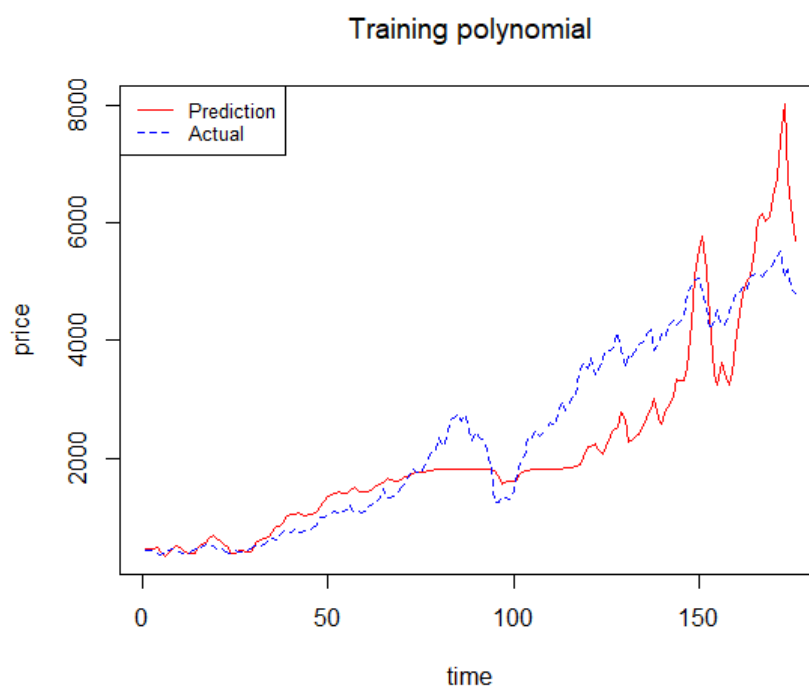
Berdasarkan gambar 2 tampak hasil prediksi dan aktual yang sangat menyerupai artinya untuk model dengan kernel linear dapat dikatakan memiliki hasil prediksi dan aktual training yang baik.



Gambar 3. *Plot Prediction and Actual Testing with Linear Kernel*

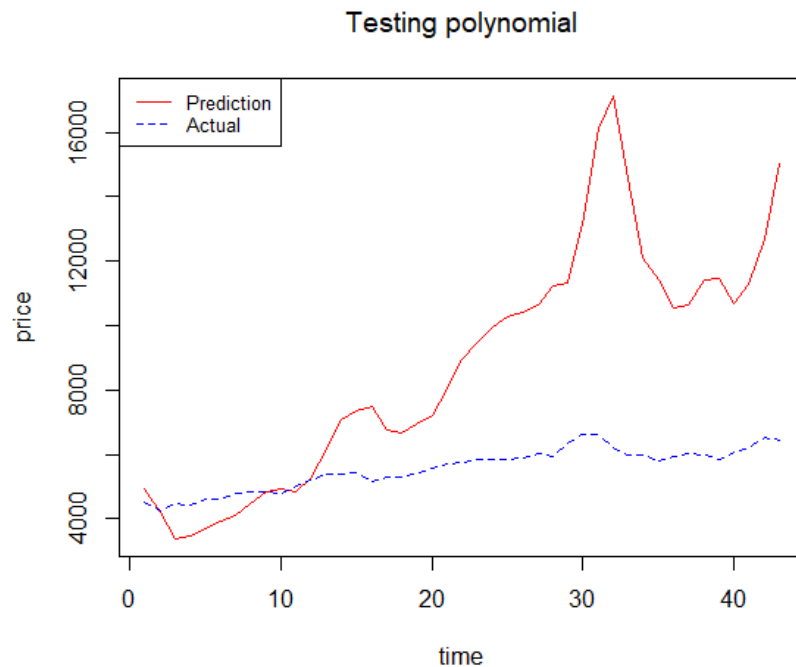
Kemudian setelah dilakukan pada data test, berdasarkan gambar 3 tampak hasil prediksi dan aktual yang sangat menyerupai artinya untuk model dengan kernel linear dapat dikatakan memiliki hasil prediksi dan aktual untuk data test baik.

3.4.2. Polynomial Kernel



Gambar 4. *Plot Prediction and Actual Training with Polynomial Kernel*

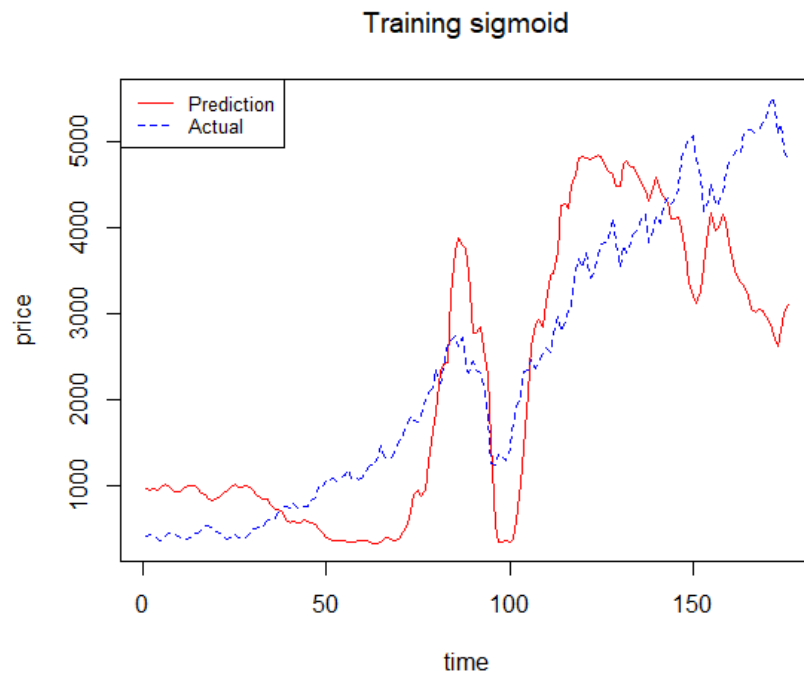
Berdasarkan gambar 4 tampak hasil prediksi dan aktual yang tidak begitu sama. Model dengan kernel polynomial dapat dikatakan memiliki hasil prediksi dan aktual training kurang baik.



Gambar 5. *Plot Prediction and Actual Test with Polynomial Kernel*

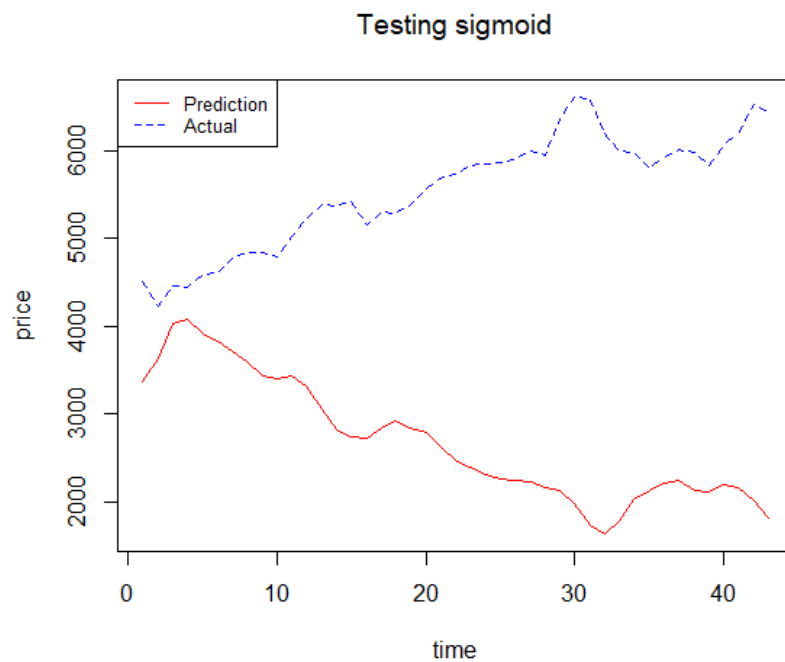
Kemudian setelah dilakukan pada data test, berdasarkan gambar 5 tampak hasil prediksi dan aktual yang cukup berbeda jauh terutama pada waktu 10 sampai ke 40 dimana nilai prediksi jauh meningkat lebih tinggi dibandingkan nilai aktual yang artinya untuk model dengan kernel polynomial dapat dikatakan memiliki hasil prediksi dan aktual untuk data test kurang baik.

3.4.3. Sigmoid Kernel



Gambar 6. *Plot Prediction and Actual Training with Sigmoid Kernel*

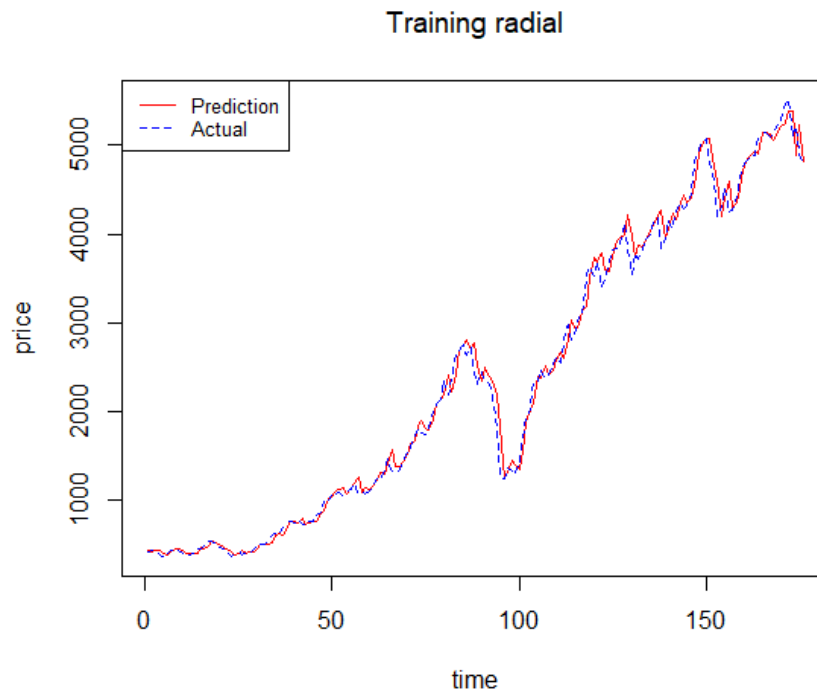
Berdasarkan gambar 6 tampak hasil prediksi dan aktual yang cukup berbeda. Dimana pada di antara waktu ke-50 menunjukkan hasil prediksi dan aktual yang berbeda dimana prediksi menurun sementara aktual naik. Hal tersebut juga terjadi pada waktu lain seperti di antara waktu ke-150.



Gambar 7. *Plot Prediction and Actual Test with Sigmoid Kernel*

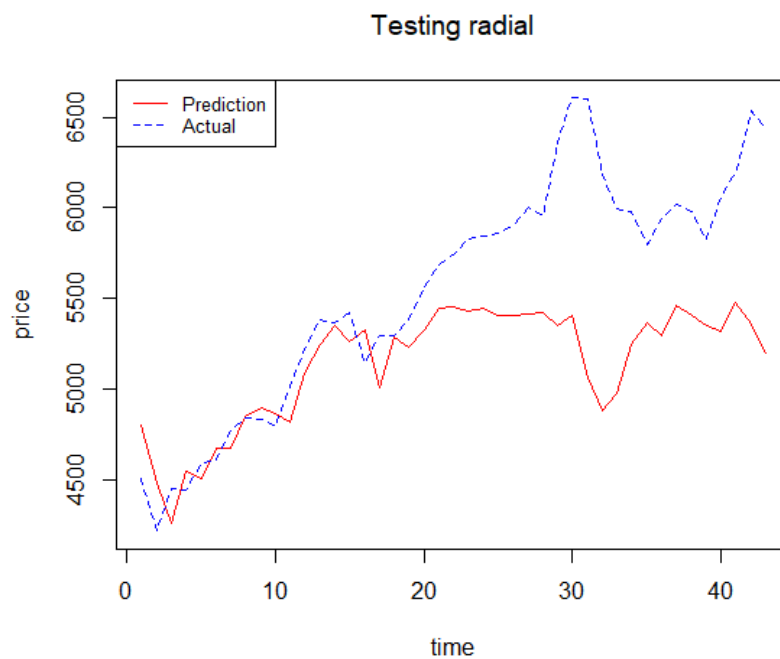
Kemudian setelah dilakukan pada data test, berdasarkan gambar 7 tampak hasil prediksi dan aktual yang sangat berbeda jauh dengan hasil yang berkebalikan dimana nilai aktual naik sementara nilai prediksi menurun selama waktu berturut-turut.

3.4.4. Radial Kernel



Gambar 8. *Plot Prediction and Actual Training with Radial Kernel*

Berdasarkan gambar 8 tampak hasil prediksi dan aktual yang sangat menyerupai artinya untuk model dengan kernel radial dapat dikatakan memiliki hasil prediksi dan aktual training yang baik.



Gambar 9. *Plot Prediction and Actual Test with Radial Kernel*

Namun setelah dilakukan pada data test tampak pada waktu ke 20 sampai 40 lebih memiliki nilai prediksi dan aktual yang cukup berbeda dimana nilai aktual melambung lebih tinggi dibandingkan nilai prediksi.

3.5. Kecocokan Model

Setelah eksplorasi hasil plot prediksi dan aktual akan dilakukan juga pemilihan model dengan melihat nilai *R-Square* dan MAPE(*Mean Absolute Percentage Error*) untuk *Training* dan *Testing* pada setiap kernel. Hasil dari *R-Square* dan MAPE sebagai berikut.

Tabel 3. Kecocokan Model

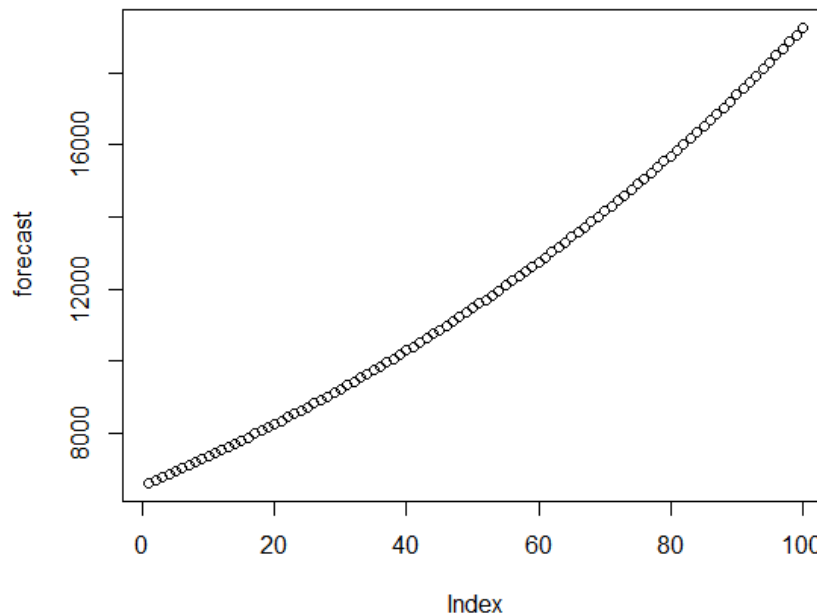
Kernel	Training		Testing	
	<i>R-Square</i>	MAPE	<i>R-Square</i>	MAPE
Linear	0,9925856	4,89308	0,9292418	2,387365
Polynomial	0,7656432	22,8834	-46,94094	57,54443
Sigmoid	0,6482374	51,60726	-23,20759	49,43487
Radial	0,9930869	4,503163	0,1152092	7,600009

Dari hasil tabel di atas tampak bahwa model yang paling cocok untuk *SVR* yaitu model kernel linear dengan nilai *R-Square* 99,25856% dan MAPE 4,89308 pada *training* dan *R-Square* 92,92418% dan MAPE 2,387365 pada *testing*. Meskipun pada *training* dapat dilihat model kernel radial memiliki nilai kecocokan model lebih baik dengan *R-Square* 99,30869% dan MAPE 4,503163 namun pada *testing* memiliki hasil kecocokan yang jauh dibandingkan dengan model kernel linear dengan *R-Square* 11,52092% dan MAPE 7,600009.

Oleh karena itu diambil model kernel linear. Kemudian pada model kernel polynomial dan model kernel sigmoid tampak memiliki nilai MAPE yang besar dibandingkan model kernel linear dan radial, bahkan memiliki *R-Square* yang negatif pada *testing* yang menunjukkan bahwa model kernel tidak cocok untuk data tersebut.

3.5. Peramalan (*Forecast*)

Dengan model yang sudah dipilih akan dilakukan peramalan untuk 100 periode kedepan. Kemudian didapatkan hasil dari peramalan selama 100 periode kedepan sebagai berikut.



Gambar 10. Hasil Peramalan selama 100 periode kedepan

Dari hasil peramalan di atas dapat diketahui bahwa harga saham JKSE akan naik dari kurang dari 8000 sampai lebih dari 16000 selama 100 periode kedepan.

4. Kesimpulan

Dengan Metode SVR (*Support Vector Regression*) didapatkan model terbaik yaitu model dengan kernel linear, *Cost* 100, dan *Epsilon*(ϵ) 0,01 dengan kecocokan model nilai *R-Square* 99,25856% dan *MAPE* 4,89308 pada *training* dan *R-Square* 92,92418% dan *MAPE* 2,387365 pada *testing*. Kemudian setelah dilakukan peramalan selama 100 periode kedepan tampak bahwa harga saham JKSE akan naik selama 100 periode kedepan. Dari hasil peramalan ini, dapat dijadikan sebagai bahan untuk dijadikan pilihan kepada pembeli saham apakah perlu membeli saham JKSE atau tidak.