



Vol. 7 No. 2 (2022) 63 - 70

JOINTECS

(Journal of Information Technology and Computer Science)

e-ISSN:2541-6448

p-ISSN:2541-3619

Support Vector Regression Dalam Prediksi Penurunan Jumlah Kasus Penderita Covid-19

Dodi Suprayogi¹, Hilman Ferdinandus Pardede²

^{1,2} Program Studi Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Nusa Mandiri

² Pusat Penelitian Informatika, Lembaga Ilmu Pengetahuan Indonesia

¹dodievangalion@gmail.com, ²hilman@nusamandiri.ac.id

Abstract

The spread of the COVID-19 virus is very worrying and continues to spread and spread throughout the infected country from children to adults. There are many ways to make a prediction, in this case determining the prediction of the number of COVID-19 sufferers can use machine learning, not only COVID-19. This study tries to predict when the COVID-19 Pandemic will decline by using the SVR algorithm with RBF, Linear, Polynomial, and Sigmoid kernels, selecting the model using SVR because SVR is able to overcome overfitting. This study uses a dataset from the Johns Hopkins University github using a sample of five countries with different numbers of COVID-19 cases. The results obtained for the RBF kernel are very good for five countries in making a graph pattern that fits between the actual data and the predicted data, by tuning different parameters in each country, then testing the gamma value to get the RMSE, R2, and MAE values, The best results are in Germany with an RMSE value of 0.099, then Italy with an RMSE value of 0.101, Indonesia with an RMSE value of 0.102, Brazil with an RMSE value of 0.105, and the US with an RMSE value of 0.105.

Keywords: COVID-19; SVR; kernel; prediction.

Abstrak

Penyebaran Virus COVID-19 sangat mengkhawatirkan dan terus menyebar serta meluas di seluruh negara terinfeksi mulai dari anak-anak hingga orang dewasa. Banyak cara dalam membuat suatu prediksi, dalam hal ini menentukan prediksi jumlah penderita COVID-19 bisa dengan menggunakan *machine learning*, tidak hanya COVID-19. Penelitian ini mencoba memprediksi kapan *Pandemic COVID-19* ini menurun dengan menggunakan algoritma SVR dengan *kernel* RBF, *Linear*, *Polynomial*, dan *Sigmoid*, pemilihan model menggunakan SVR karena SVR mampu mengatasi *overfitting*. Penelitian ini menggunakan dataset dari *github John Hopkins University* menggunakan sample lima negara dengan jumlah kasus COVID-19 yang berbeda. Hasil yang didapat untuk *kernel* RBF sangat baik untuk lima negara dalam membuat pola grafik yang fit antara data aktual dan data prediksi, dengan melakukan *tuning parameter* yang berbeda-beda di setiap negara, kemudian melakukan pengujian nilai gamma untuk mendapatkan nilai RMSE, R2, dan MAE, hasil terbaik ada pada negara Jerman dengan nilai RMSE 0.099, kemudian Itali dengan nilai RMSE 0.101, Indonesia nilai RMSE 0.102, brazil nilai RMSE 0.105, dan US nilai RMSE 0.105.

Kata kunci: COVID-19; SVR; kernel; prediksi.



1. Pendahuluan

Pandemic COVID-19 sangat mengkhawatirkan masyarakat hampir di seluruh dunia setiap harinya jumlah penderita di beberapa negara mengalami kenaikan yang signifikan. Bagi masyarakat dan khususnya dunia usaha perlu adanya informasi prediksi kapan COVID-19 bisa segera berakhir hal itu merupakan

data penting yang dapat dijadikan bahan untuk membuat rencana. Tanpa ada informasi kapan wabah Covid-19 menurun, tentu akan membuat keadaan menjadi serba tidak menentu. Dalam ketidakpastian, rencana tidak dapat disusun. Dan akibatnya, dunia usaha, mengalami kemerosotan drastis, sebagian mungkin telah gulung tikar. Prediksi seperti angin pembawa harapan bagi masyarakat untuk kembali hidup normal, dan

Diterima Redaksi : 22-04-2022 | Selesai Revisi : 30-05-2022 | Diterbitkan Online : 31-05-2022

beraktifitas seperti sedia kala. Banyak cara dalam membuat suatu prediksi, dalam hal ini menentukan prediksi jumlah penderita COVID-19 bisa dengan menggunakan machine learning, tidak hanya COVID-19, prediksi yang dibuat dengan machine learning mampu memprediksi seperti saham, ekonomi, dan permintaan. Sehingga sangat diperlukan informasi prediksi dalam memudahkan masyarakat untuk dapat menerima informasi contohnya tentang COVID-19. Beberapa penelitian yang sudah dilakukan mengenai COVID-19.

Penelitian pertama menggunakan metode Naïve Bayes diperoleh prediksi positif yang menunjukkan angka 55,48% dan negatif hasil prediksi 44,52%. Penelitian pertama menghasilkan klasifikasi data yang telah diperoleh bahwa prediksi terbesar ditemukan pada status positif terinfeksi virus COVID-19 mencapai 55,48%. Penelitian pertama ini menghasilkan hasil yang baik dan dapat digunakan sebagai referensi untuk masyarakat agar dapat mengetahui klasifikasi status terinfeksi COVID-19[1].

Penelitian kedua menggunakan beberapa model yaitu linear regresi, LASSO regresi, dan *Exponential Smoothing*. Penelitian ini mengambil *sample* dari jumlah kematian dan jumlah yang pulih dari COVID-19 berdasarkan *time series* dari tahun 2019, Sistem menganalisa kumpulan data yang berisi data masa lalu yang aktual dan membuat prediksi yang akan datang, hasil studi membuktikan bahwa model *Exponential Smoothing* mampu bekerja lebih baik dalam memprediksi tingkat kematian dan mengkonfirmasi kasus COVID-19, menurut hasil dari model linear regresi dan LASSO regresi tingkat kematian akan meningkat di masa mendatang dan tingkat pemulihan yang lambat.[2]. Konsep SVR didasarkan pada risk minimization, yaitu untuk mengestimasi suatu fungsi dengan cara meminimalkan batas atas dari generalization error, sehingga SVR mampu mengatasi overfitting dan SVR yang lebih rendah dari metode seperti neural network selain itu Support Vector Machine menggunakan *kernel* untuk mentransformasikan ruang input ke feature space atau mengimplementasikan model ke dimensi yang lebih tinggi sehingga kasus non *Lineary* separable pada ruang input menjadi *Lineary* separable pada feature space fungsi *kernel* ini memiliki banyak jenis diantaranya *kernel Radial Basis Function (RBF)*, *kernel Linear*, dan *kernel Polynomial* [7]. Dan dari beberapa *kernel* tersebut beberapa perbedaan ketiga fungsi *kernel* tersebut terdapat pada fungsi pemetaannya ke dalam feature space, dan masing-masing *kernel* memiliki kelebihan dan kekurangan yaitu pada setiap kasus tertentu sehingga perlu diadakannya percobaan dalam mencari fungsi *kernel* terbaik.

Tujuan penelitian ini untuk memprediksi penurunan jumlah penderita pandemic COVID-19 dengan menggunakan metode SVR dengan *Kernel RBF*,

Sigmoid, *Polynomial*, dan *Linear Regression*. Dengan mengambil *sample* lima negara dengan kasus yang berbeda. Dan melihat seberapa akurat metode dengan *kernel* tersebut dalam membentuk prediksi. Hasil yang akan diberikan oleh SVR akan berbentuk kurva yang fit antara data aktual dengan data prediksi serta memprediksi jumlah penderita COVID-19 untuk hari-hari berikutnya.

2. Metode Penelitian

SVR merupakan penerapan support vector machine (SVM) untuk kasus regresi. Dalam kasus regresi output berupa bilangan *riil* atau *continue*. SVR merupakan metode yang dapat mengatasi overfitting, sehingga akan menghasilkan performansi yang bagus. Model SVR sering digunakan untuk meminimalkan jumlah *square error* Algoritme SVR adalah teori yang diadaptasi dari teori *machine learning* yang sudah digunakan untuk memecahkan masalah klasifikasi.

2.1. Dataset

Pada Tabel 1 ini dataset yang digunakan adalah data yang didapat dari *github John Hopkins University*. Data berisikan 165 negara yang terinfeksi COVID-19, adapun atribut dari 165 negara seperti propinsi, negara, koordinat dan long serta tanggal. Dataset ini yang digunakan untuk penelitian menggunakan model SVR, dataset ini akan dibersihkan Kembali jika ditemukan data yang tidak valid / kosong.

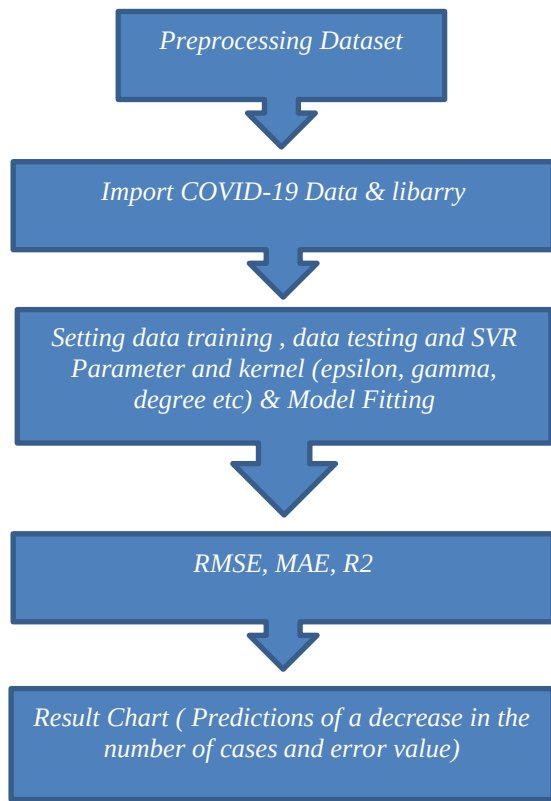
Tabel 1. Deskripsi Dataset

Attribute	Description
Province/State	Name Of Province
Country/Region	Name of Country
Lat	Country Latitude
Long	Country Longitude
Date	Date of addition case infection COVID-19

Dataset yang dikumpulkan biasanya tidak langsung bisa diolah sehingga butuh dikoreksi seperti kolom yang tidak sesuai, seperti dataset yang didapat dari John Hopkins dataset tersebut tidak memiliki kolom tanggal sehingga perlu dibuat kolom tanggal agar memudahkan pada saat melakukan prediksi. Adapun beberapa dataset ada yang selalu memerlukan update seperti kasus penambahan penderita COVID-19, karena setiap hari jumlah penderita bertambah maka perlu dilakukan penambahan dataset secara berkala agar prediksi yang dilakukan bisa menghasilkan prediksi yang baik. Dataset yang baik bisa didapat dari membuat prototype sendiri lalu di uji coba dan akan menghasilkan data yang baik dan akurat sehingga data tersebut bisa diolah untuk penelitian lebih lanjut, ada juga dataset yang memang dikumpulkan oleh peneliti agar bisa digunakan bersama untuk penelitian masing-masing. Tahapan penelitian yang digunakan dalam penelitian ini dimulai dari Preprocessing dataset, Import COVID-19 Data & Library, Setting data training, data testing and SVR

Parameter and kernel (*epsilon, gamma, degree etc*) & Model Fitting, terlihat pada Gambar 1.

2.2. Tahapan Penelitian



Gambar 1. Tahapan Penelitian

Berikut ini merupakan penjelasan dari tahapan penelitian yang terdapat pada Gambar 1. *Dataset* yang diambil dari *github John Hopkins University* pada tahap *preprocessing* fokus utama yang dilakukan adalah memproses data mentah untuk menghasilkan data yang siap untuk membentuk model prediksi. Hal pertama yang dilakukan adalah memeriksa apakah ada nilai yang hilang pada *dataset*. Tahapan *preprocessing* data berfungsi untuk membantu memperbaiki kinerja algoritma, sehingga tingkat keberhasilannya lebih mendekati nilai sebenarnya. Setelah selesai tahapan *preprocessing*, kemudian dilanjutkan melakukan import *dataset* dan *library* seperti *SVR, R2_score, split test* pada *machine learning* dan *machine learning* yang digunakan adalah *python*. Setelah *dataset* dan *library* di import kemudian *dataset* dipisahkan menjadi *dataset training* dan *dataset testing* lalu diuji dengan menggunakan metode *SVR (Support Vector Regression)* dengan beberapa *Kernel* yaitu *RBF, Sigmoid, Linear, Polynomial* kernel tersebut harus menentukan atau melakukan *tunning parameter* disetiap kernel agar dalam pengujian bisa menghasilkan hasil yang baik, seperti *gamma, epsilon, degree*, dan lainnya pada masing-masing *Kernel* hingga mendapatkan hasil yang maksimal. Setelah diuji untuk evaluasi pengujian akan diperoleh nilai seperti *Root Mean Square Error (RMSE)*, *Mean Absolute Error (MAE)* dan *R-Square*

(*R2*). Pada program jurnal linear jika ingin digunakan pada kernel *RBF, polynomial, sigmoid* cukup diganti pada tulisan '*linear*' dengan kernel yang ingin dipakai sedangkan *tunning* parameternya disesuaikan karena akan mempengaruhi hasil yang diinginkan, untuk *RMS, MAE, R2* dimana data training dan data prediksi dilakukan perhitungan sesuai rumus yang dimasukkan.

Program Jurnal

```
svr=SVR(kernel='linear', C=1e3, gamma='scale',
max_iter=-1, cache_size=200, tol=1)
```

```
svr.fit(x_train, y_train)
```

```
x_test=x_train.reshape(-1,1)
```

```
y_pred=svr.predict(x_test)
```

```
mae = mean_absolute_error(y_train, y_pred)
```

```
rms= sqrt(mean_squared_error(y_train, y_pred))
```

```
r2 = r2_score(y_train,y_pred)
```

Setelah melakukan perbandingan akan didapat hasil yang terbaik berupa grafik yang fit antara data aktual dan data prediksi kemudian akan tampak hasil penurunan jumlah kasus COVID-19 di setiap negara. Sehingga hasil yang akan didapat tergantung dari ketepatan dalam mengatur nilai parameter pada masing-masing *kernel*. Adapun nilai default parameter terlihat pada Tabel 2 [18], nilai tersebut adalah nilai awal ketika ingin menggunakan beberapa kernel pada *SVR*, yaitu *RBF* dengan *C* adalah nilai parameter regulasi, *Gamma* adalah kernel koefisien untuk *rbf, poly*, dan *sigmoid* nilai awal adalah *scale*, *Epsilon* yaitu sembarang nilai positif (nilai yang paling kecil atau bisa yang terbesar) dimana nilai positif inilah yang membantu untuk menaksir nilai limit dan kebenaran nilainya, *Coef0* adalah independen dalam fungsi kernel hanya digunakan pada kernel *polynomial* dan *sigmoid* nilai defaultnya adalah 0, *Degree* adalah fungsi kernel yang digunakan pada *polynomial* dan Suku pangkat tertinggi dalam *polynomial* nilai default pada *polynomial* adalah 3.

Tabel 2. Nilai Default Parameter SVR

Kernel	C	Gamma	Epsilon	Coef0	Degree
RBF	1	Scale	0.1		
Linear	1		0.1		
Polynomial	1	Scale	0.1	0	3
Sigmoid	1	Scale	0.1	0	

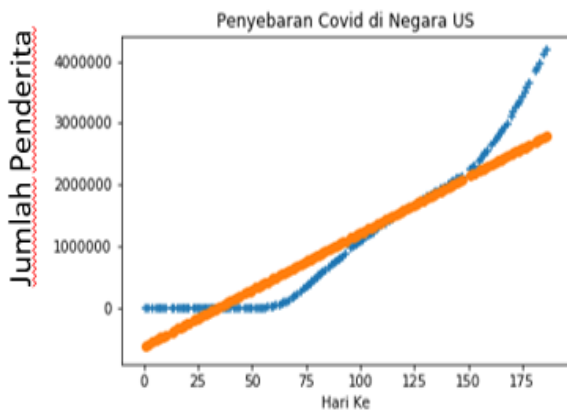
2.3 Parameter Tuning

Parameter tuning dilakukan untuk mencari nilai *parameter* paling optimal dalam membentuk model prediksi terbaik. Proses *parameter tuning* berfokus pada pemilihan nilai parameter *kernel* seperti *gamma, C, coef0, degree, epsilon, max_iter, Tol*, dan *Cache_size*. Pengujian untuk menentukan nilai *parameter* yang tepat perlu dilakukan secara berulang [16]. Setelah mendapatkan parameter yang paling optimal, yaitu dengan akurasi model terbaik, model kembali diuji untuk dievaluasi menggunakan data uji.

3. Hasil dan Pembahasan

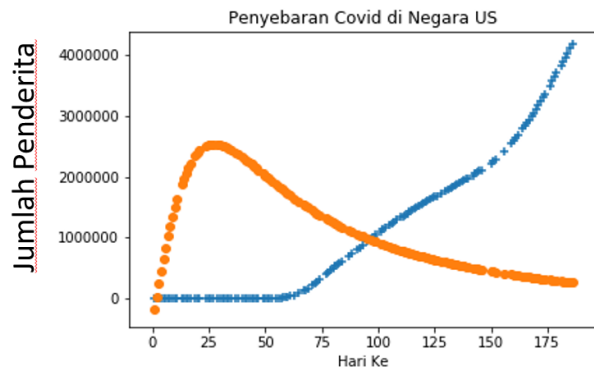
Dalam melihat hasil suatu prediksi dilakukan suatu evaluasi. Evaluasi tersebut digunakan untuk mengetahui keakuratan hasil peramalan yang telah dilakukan terhadap data yang sebenarnya. Beberapa metode dapat digunakan untuk melakukan perhitungan kesalahan peramalan dan metode yang digunakan diantaranya adalah *Root Mean Squared Error (RMSE)*. RMSE merupakan metode yang cukup sering digunakan dalam mengevaluasi kinerja prediksi dengan menggunakan RMSE. Hal yang membuat berbeda karena adanya keacakan pada data atau karena tidak mengandung estimasi yang akurat. RMSE merupakan akar nilai dari MSE yang sudah dicari sebelumnya, RMSE digunakan untuk mencari nilai akurasi hasil peramalan dengan data history dengan menggunakan rumus yang ditunjukkan dengan nilai kesalahan. Semakin kecil nilai yang dihasilkan semakin bagus pula hasil peramalan yang dilakukan [17].

Eksperimen kesatu negara US menggunakan semua *kernel*. Pada Pengujian pertama dilakukan pada negara US, dimana jumlah penderita dan tanggal atau hari ke, sebagai nilai x dan y, pengujian dilakukan menggunakan SVR dengan *Kernel Linear, Sigmoid, Polynomial, dan RBF*. Hasil dari pengujian bervariasi dimulai dari hasil yang sangat jauh antara data aktual dengan data prediksi, hingga hasil prediksi yang sangat terlihat fit dengan data aktual. Hasil COVID-19 yang bervariasi tersebut menggambarkan model yang cocok digunakan untuk memprediksi kasus jumlah penderita COVID-19



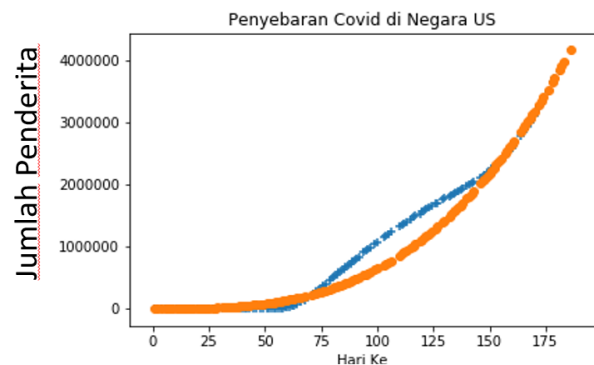
Gambar 2. Kernel Linear Negara US

Terlihat garis warna biru untuk data aktual dengan jumlah penderita terus bergerak naik dan garis warna kuning untuk data prediksi. Pada Gambar 2 *kernel* yang digunakan adalah *Linear* hasil yang ditampilkan Gambar 2 tidak fit antara data aktual dengan data prediksi. Jumlah penderita semakin naik sedangkan hasil prediksi tidak bisa mengikuti hasil data aktual dan tidak membentuk kurva yang fit, maka untuk kernel linear tidak bisa memprediksi penderita COVID-19 dengan baik. Kernel linear membentuk garis lurus tidak bisa membentuk garis kurva yang fleksibel sehingga linear, walaupun parameter di *tunning* dengan baik tetap kurva yang dibentuk linear tidak bisa mengikuti data aktual.



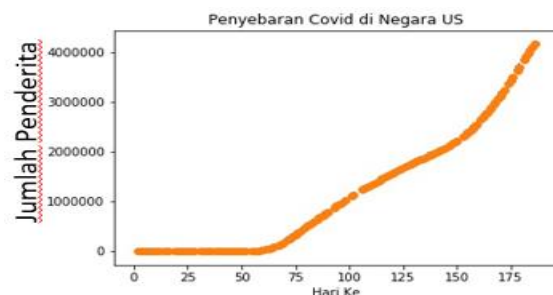
Gambar 3. Kernel Sigmoid Negara US

Kemudian dilanjutkan dengan pengujian menggunakan *Kernel Sigmoid*, pada Gambar 3 terlihat kurva *Sigmoid* tidak bisa memprediksi penderita COVID-19. Terlihat data aktual jumlah penderita COVID-19 tidak bisa fit dengan data prediksi, kurva yang dibentukpun oleh kernel sigmoid tidak fit. Sehingga kernel sigmoid tidak bisa untuk melakukan prediksi terhadap jumlah penderita COVID-19.



Gambar 4. Kernel Polynomial Negara US

Lalu terlihat pada Gambar 4 menggunakan *Kernel Polynomial* untuk jumlah penderita COVID-19 mulai terlihat fit antara data prediksi dengan data aktual. Tetapi kurva prediksi jumlah penderita COVID-19 tidak seluruhnya fit masih yang meleset dari data aktual. Sehingga metode polynomial tidak bisa memprediksi dengan baik jumlah penderita COVID-19 yang akan datang.



Gambar 5. Kernel RBF Negara US

Terakhir terlihat pada Gambar 5 kurva yang ditampilkan tampak fit antara data prediksi dan aktual untuk memprediksi jumlah penurunan penderita COVID-19.

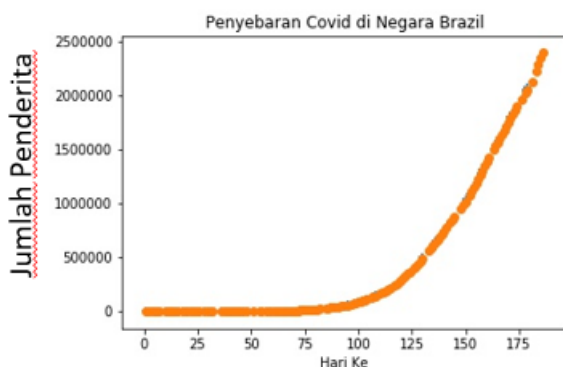
Untuk prediksi terlihat adanya jumlah penderita COVID-19 yang sangat fit sehingga untuk *Kernel RBF* terlihat lebih bisa memprediksi jumlah penderita COVID-19. Nilai *parameter* pada kernel RBF yang digunakan dari sample untuk memprediksi jumlah penderita COVID-19 ada pada Tabel 3 :

Tabel 3. Nilai Parameter pada Negara US

Kernel	C	Gamma	Epsilon	Coef0	Degree
Linear	10	Scale			
Polynomial	10	0.1	0.1		2
Sigmoid	99000	Scale	0.1	0.1	
RBF	600000	0.1	0.1		

Pengujian untuk memprediksi jumlah penderita COVID-19 dilakukan berulang pada kernel Linear, Polynomial, Sigmoid, dan RBF seperti merubah nilai C pada setiap *kernel* atau gamma agar menghasilkan kurva yang fit antara data aktual dengan data prediksi. Pengujian pada data jumlah penderita COVID-19 yang sama pun dilakukan terhadap beberapa negara lainnya, dengan menggunakan teknik yang sama hanya nilai pada parameter saja yang berubah. Untuk nilai parameter yang dirubah dalam memprediksi jumlah penderita COVID-19 pada negara US dengan gamma 0.25 menjadi yang paling baik dengan nilai RMSE 0.105 dengan data prediksi 4178020.880 mendekati data aktual, dengan nilai RMSE yang mendekati angka 0 maka kurva yang dibentuk dipastikan sangat fit.

Eksperimen kedua negara Brazil menggunakan semua *Kernel*, percobaan yang dilakukan terhadap negara Brazil untuk *kernel Linear, Polynomial, Sigmoid* sama dengan negara US tidak bisa fit antara data aktual dengan data prediksi sehingga percobaan yang paling baik dengan menggunakan kernel RBF. Kernel RBF dalam membentuk kurva yang fit pada jumlah penderita COVID-19 terlihat sangat baik antara data aktual dengan data prediksi terlihat *Kernel RBF* pada Gambar 6. Kurva yang terbentuk oleh data aktual dan data prediksi mampu membentuk garis yang fit, kurva tersebut terlihat makin menunjukkan jumlah penderita yang bertambah setiap harinya.



Gambar 6. Kernel RBF Negara Brazil

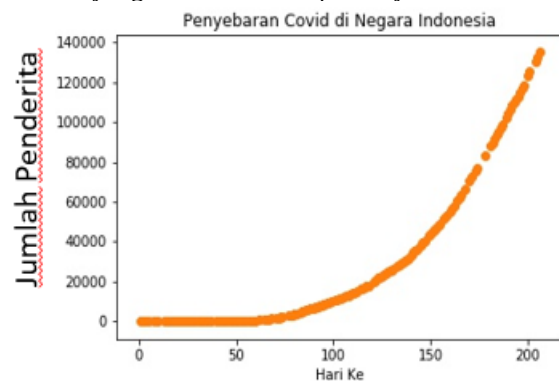
Pada tabel 4 terlihat nilai parameter pada *kernel RBF*, pengujian pun dilakukan berulang untuk mendapatkan hasil yang maksimal pada *kernel RBF*. Sedangkan untuk nilai parameter lainnya berusaha di *tunning* maksimal

tetapi masih belum fit antara data aktual dan data prediksi untuk jumlah penderita COVID-19. Tabel 4 pada pengujian negara Brazil untuk gamma 1 mendapatkan nilai RMSE terkecil yaitu 0.10501 dengan nilai data prediksi 2394512.891, nilai RMSE yang mendekati angka 0 ada pada RBF gamma 1.

Tabel 4. Nilai Parameter pada Negara Brazil

Kernel	C	Gamma	Epsilon	Coef0	Degree
Linear	10	Scale			
Polynomial	10	0.1	0.1		2
Sigmoid	70000	Scale	0.1	0.1	
RBF	600000	1	0.1		

Eksperimen ketiga negara Indonesia menggunakan semua *kernel* agar dapat diketahui pasti *kernel yang baik*. Kernel RBF dalam membentuk kurva yang fit pada jumlah penderita COVID-19 terlihat sangat baik antara data aktual dengan data prediksi terlihat Kernel RBF pada Gambar 7. Kurva yang terbentuk oleh data aktual dan data prediksi mampu membentuk garis yang fit, kurva tersebut terlihat makin menunjukkan jumlah penderita yang bertambah setiap harinya.



Gambar 7. Kernel RBF Negara Indonesia

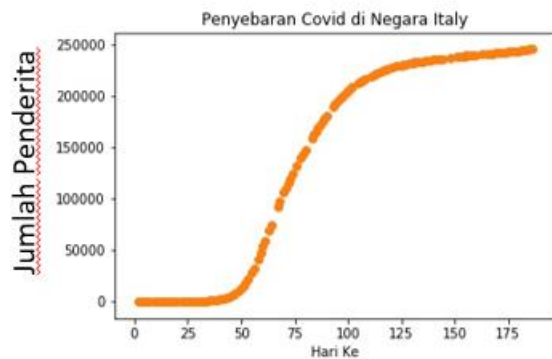
Pada Tabel 5 terlihat nilai parameter pada kernel RBF, pengujian pun dilakukan berulang untuk mendapatkan hasil yang maksimal pada kernel RBF. Sedangkan untuk nilai parameter lainnya berusaha di *tunning* maksimal tetapi belum berhasil masih belum fit antara data aktual dan data prediksi untuk jumlah penderita COVID-19. Tabel 5 pada pengujian negara Indonesia untuk gamma 0.15 dengan RMSE 0.1021, nilai data prediksi 1385634.900 nilai parameter tersebut harus sangat disesuaikan dan tepat agar hasil kurvanya maksimal.

Tabel 5. Nilai Parameter Pada Negara Indonesia

Kernel	C	Gamma	Epsilon	Coef0	Degree
Linear	10	Scale			
Polynomial	10	0.1	0.1		2
Sigmoid	30000	Scale	0.1	0.1	
RBF	2000000	0.1	0.1		

Eksperimen keempat negara Itali menggunakan semua *Kernel* percobaan yang dilakukan terhadap negara Itali dengan beberapa percobaan nilai parameter. Untuk *kernel Linear, Polynomial, Sigmoid* sama dengan negara US tidak bisa fit antara data aktual dengan data prediksi sehingga percobaan yang paling baik dengan

menggunakan kernel RBF. Kernel RBF pada negara Itali membentuk kurva yang fit pada jumlah penderita COVID-19 terlihat sangat baik antara data aktual dengan data prediksi terlihat *Kernel RBF* pada Gambar 8.

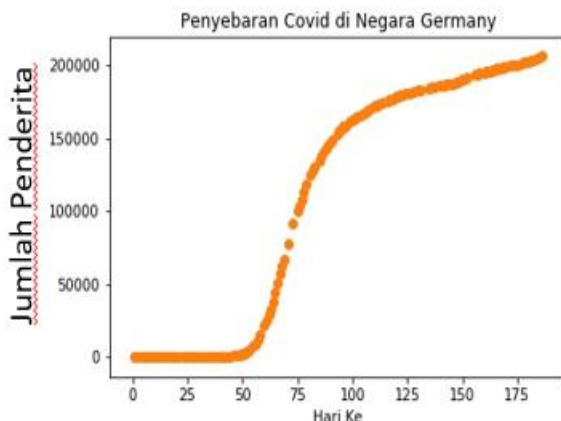


Gambar 8. *Kernel RBF* Negara Italy

terlihat nilai parameter pada kernel RBF, pengujian pun dilakukan berulang untuk mendapatkan hasil yang maksimal pada kernel RBF. Sedangkan untuk nilai parameter lainnya berusaha di *tunning* maksimal tetapi masih belum berhasil fit antara data aktual dan data prediksi untuk jumlah penderita COVID-19. Tabel 6 pada pengujian negara Itali untuk gamma 0.1 dengan RMSE 0.1021, nilai parameter tersebut harus sangat disesuaikan dan tepat agar hasil kurvanya maksimal.

Tabel 6. Nilai Parameter pada Negara Italy

Kernel	C	Gamma	Epsilon	Coef0	Degree
Linear	10	Scale			
Polynomial	10	0.1	0.1		2
Sigmoid	300000	Scale	0.1	0.1	
RBF	600000	0.1	0.01		



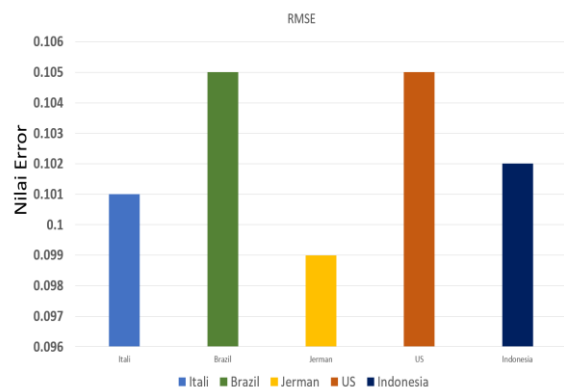
Gambar 9. *Kernel RBF* Negara Jerman

Pada pengujian kelima Gambar 9, mendapatkan hasil baik karena terlihat sangat fit antara data aktual dengan data prediksi. Hasil yang didapat untuk Negara Jerman dengan gamma 1 adalah yang paling baik mendapatkan nilai RMSE 0.100732 dan nilai prediksi 206277.899 mendekati data aktual. *Tunning* parameter pada negara Jerman dilakukan secara berulang agar menghasilkan nilai *error* yang kecil, nilai parameter terlihat pada Tabel 7 dimana nilai tersebut digunakan pada metode RBF.

Tabel 7. Nilai Parameter pada Negara Jerman

Kernel	C	Gamma	Epsilon	Coef0	Degree
Linear	10	Scale			
Polynomial	10	0.1	0.1		2
Sigmoid	300000	Scale	0.1	0.1	
RBF	900000	0.1	0.01		

Hasil analisa pada seluruh pengujian bahwa untuk membuat kurva yang akurat antara data aktual dengan data prediksi hanya bisa dilakukan oleh *kernel RBF*, jika dilihat dari kurva di setiap negara yang dibentuk oleh kurva tersebut maka terlihat jelas kurva membentuk data yang fit, dibandingkan dengan *kernel* lainnya walaupun parameter dirubah dengan sebaik mungkin tetap saja hanya *kernel RBF* yang paling baik. Begitu juga dengan hasil *error* yang ditampilkan seperti RMSE untuk *kernel RBF* terlihat paling mendekati angka 0, jika nilai RMSE mendekati angka 0 maka prediksi yang dihasilkan pun menjadi lebih baik seperti contoh pada negara US nilai RMSE 0.105, selain itu untuk mendapatkan nilai *error* yang kecil maka nilai parameter harus *tunning* terlebih dahulu seperti nilai gamma, C, coef, dan degree. Dengan demikian jika *kernel* menghasilkan kurva yang fit lalu nilai *error* kecil maka akan menghasilkan prediksi yang baik terlihat pada Gambar 10 adalah hasil perbandingan nilai RMSE pada setiap negara terlihat bahwa negara Jerman mempunyai nilai RMSE 0.99 kemudian nilai RMSE pada negara Itali 0.101, setelah negara Itali dilanjut negara Indonesia dengan nilai RMSE 0.102, untuk negara Brazil dengan nilai RMSE 0.105, nilai RMSE Brazil sama dengan nilai RMSE yang ada pada negara US.

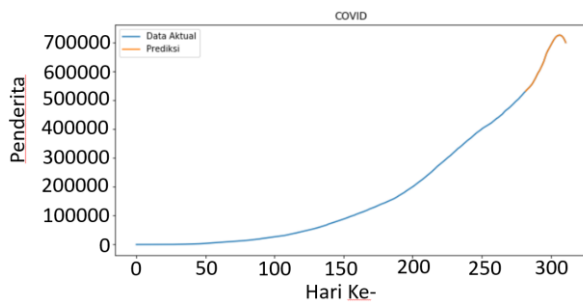


Gambar 10. Hasil Perbandingan Nilai RMSE Setiap Negara

3.1. Hasil Prediksi

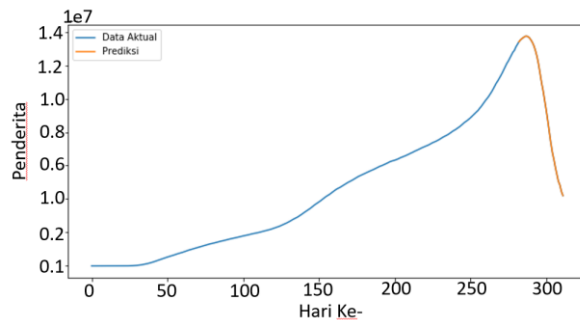
Setelah melakukan beberapa kali pengujian pada beberapa negara untuk melihat penurunan jumlah penderita COVID-19, hasil yang diperoleh setiap negara berbeda. Prediksi yang dihasilkan dilakukan pengujian dengan *kernel RBF* dimana *kernel* tersebut menghasilkan kurva yang sangat fit pada kurva antara data aktual dengan data prediksi dalam memprediksi penurunan jumlah penderita COVID-19 yang ada di beberapa negara. Hasil yang diujikan oleh metode SVR terlihat pada Gambar 11 prediksi COVID-19 untuk negara Indonesia mengalami penurunan jumlah

penderita COVID-19 yang belum terlalu signifikan masih sedikit berkurangnya penderita.



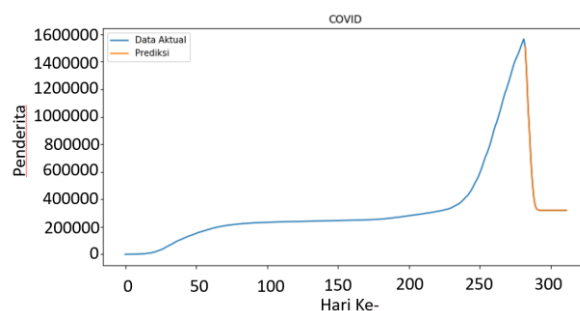
Gambar 11. Prediksi Negara Indonesia

Gambar 12 adalah hasil prediksi yang terlihat pada negara US pada jumlah penderita COVID-19. Kurva pada gambar mengalami penurunan yang signifikan pada jumlah penderita COVID-19. Sempat terlihat puncak jumlah penderita COVID-19 tetapi kemudian perlahan kurva tersebut melandai atau berkurangnya jumlah penderita COVID-19.



Gambar 12. Prediksi Negara US

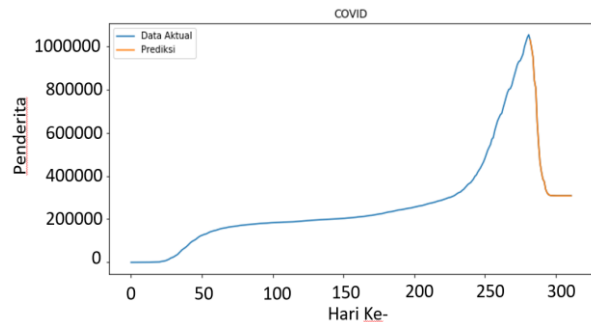
Gambar 13 adalah hasil prediksi yang terlihat pada negara Itali pada jumlah penderita COVID-19. Kurva pada gambar mengalami penurunan yang signifikan pada jumlah penderita COVID-19. Sempat terlihat puncak jumlah penderita COVID-19 kemudian perlahan kurva tersebut melandai atau berkurangnya jumlah penderita COVID-19 tetapi terlihat kurva kembali stabil setelah berkurangnya jumlah penderita COVID-19.



Gambar 13. Prediksi Negara Italy

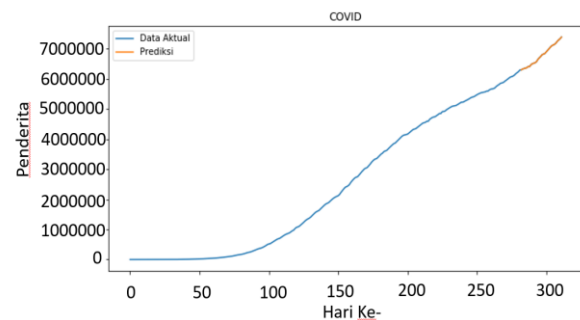
Gambar 14 adalah hasil prediksi yang terlihat pada negara Jerman pada jumlah penderita COVID-19. Kurva pada gambar mengalami penurunan yang signifikan pada jumlah penderita COVID-19. Sempat terlihat puncak jumlah penderita COVID-19 kemudian perlahan

kurva tersebut melandai atau berkurangnya jumlah penderita COVID-19 tetapi terlihat kurva kembali stabil setelah berkurangnya jumlah penderita COVID-19. Kurva pada Gambar 14 dihasilkan oleh *kernel* RBF pada saat memprediksi jumlah penurunan penderita COVID-19.



Gambar 14. Prediksi Negara Jerman

Gambar 15 adalah hasil prediksi yang terlihat pada negara Brazil pada jumlah penderita COVID-19. Kurva pada gambar mengalami kenaikan yang signifikan pada jumlah penderita COVID-19. Data prediksi tidak terlihat adanya penurunan jumlah penderita COVID-19 sehingga kurva terus mengalami kenaikan pada jumlah penderita COVID-19. Kurva pada Gambar 15 dihasilkan oleh *kernel* RBF pada saat memprediksi jumlah penurunan penderita COVID-19.



Gambar 15. Prediksi Negara Brazil

4. Kesimpulan

Dari selesainya penelitian dengan judul “*Support Vector Regression* (SVR) Dalam Prediksi Kasus Jumlah Penderita COVID-19” dapat ditarik kesimpulan sebagai berikut: Algoritma SVR dengan *Kernel Linear*, *Polynomial*, *Sigmoid* tidak dapat digunakan untuk prediksi kasus COVID-19 karena hasil yang tidak baik. Sedangkan untuk *Kernel RBF* sangat bisa digunakan untuk prediksi jumlah kasus COVID-19 karena membentuk grafik yang fit antara data aktual dengan data prediksi. Pencarian nilai RMSE pada *Kernel RBF* di setiap negara, membantu untuk mengevaluasi kinerja *Kernel* tersebut dalam memprediksi kasus COVID-19. Perubahan yang dilakukan pada parameter disetiap *kernel* sangat berpengaruh terhadap kurva yang dibuat masing-masing *kernel*. Ketika parameter tidak di *tunning* dengan baik, maka hasil kurva yang dibentuk tidak akan fit dengan data aktual. Sangat disarankan menggunakan nilai default untuk beberapa parameter.

Hasil evaluasi dan validasi yang dilakukan dengan mencari nilai Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), terkecil dan nilai R2 score yang mendekati 1, *kernel* yang diujikan yang terbaik adalah RBF dengan menggunakan *tunning* parameter gamma 1 dan 0.25 dengan hasil RMSE yang mendekati angka 0 dan R2 mendekati angka 1. Hasil nilai perbandingan RMSE pada negara yang paling rendah adalah Jerman dengan nilai 0.99 kemudian Itali 0.101, lalu Indonesia dengan nilai 0.102, Brazil dengan nilai 0.105 dan US dengan nilai 0.105.

Daftar Pustaka

- [1] Aditia, A. (2021). Covid-19: Epidemiologi, Virologi, Penularan, Gejala Klinis, Diagnosa, Tatalaksana, Faktor Risiko dan Pencegahan. *Jurnal Penelitian Perawat Profesional*, 3(4), 653-660.
- [2] Arfan, A., & ETP, L. (2020). Perbandingan Algoritma Long Short-Term Memory dengan SVR Pada Prediksi Harga Saham di Indonesia. *PETIR*, 13(1), 33 - 43.
- [3] Cahyono, R. E., Sugiono, J. P., & Tjandra, S. (2019). Analisis Kinerja Metode Support Vector Regression (SVR) dalam Memprediksi Indeks Harga Konsumen. *JTIM : Jurnal Teknologi Informasi Dan Multimedia*, 1(2), 106-116. <https://doi.org/10.35746/jtim.v1i2.22>.
- [4] Elfi Quyuni R., Moh Alimansur., 2020. Upaya Pencegahan dengan Kepatuhan dalam Pencegahan Penularan Covid-19 pada Relawan Covid. *Jurnal Recode*. No. 4. Vol. 1..
- [5] F. Rustam et al, 2020. COVID-19 Future Forecasting Using Supervised Machine Learning Models, in *IEEE Access*, vol. 8, pp. 101489-101499
- [6] Hairunisa N, Amalia H. Review: Penyakit Virus Corona Baru 2019 (COVID-19). *J Biomedika Kesehat [Internet]*. 2020 Jun. 30.
- [7] Herlawati, 2020. COVID-19 Spread Pattern Using Support Vector Regression. *Journal Penelitian Ilmu Komputer System Embedded & Logic*, Vol (1): 67 – 7
- [8] Hendayanti, Ni & Suniantara, I Ketut & Nurhidayati, Maulida. (2019). Penerapan Support Vector Regression (Svr) Dalam Memprediksi Jumlah Kunjungan Wisatawan Domestik Ke Bali. *Jurnal Varian*. 3. 43-50. 10.30812/varian.v3i1.506.
- [9] Shiyammurti, N. R., Saputri, D. A., & Syafira, E.(2020). Dampak Pandemi Covid-19 di PT. Bursa Efek Indonesia (BEI). *Journal of Accounting Taxing and Auditing (JATA)* ISSN, 1(1).
- [10] Lestari, M. I., & Anggraeni, D. (2021). Analisis Dampak Sentimen Masyarakat Selama Pandemi Covid-19 Terhadap Kurs Rupiah (Studi Kasus Pandemi Covid-19 di Indonesia). *Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi*, 9(1).
- [11] N. P. Hendayanti, I. K. P. Suniantara, and M. Nurhidayati, “Penerapan Support Vector Regression (SVR) Dalam Memprediksi Jumlah Kunjungan Wisatawan Domestik Ke Bali”, *Jurnal Varian*, vol. 3, no. 1, pp. 43-50, Oct. 2019.
- [12] Noval Dini Maulana, et al. 2019. Implementasi Metode Support Vector Regression (SVR) Dalam Peramalan Penjualan Roti (Studi Kasus: Harum Bakery) *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. Vol. 3, No. 3, hlm. 2986-2995.
- [13] Puji Hari Santoso, 2020. Application Of Data Mining Classification For Covid-19 Infected Status Using Algoritma Naïve Method, Volume 4 Number 1 May 2020, pp. 267-275
- [14] Saputra, G. H., Wigena, A. H., & Sartono, B. (2019). Penggunaan Support Vector Regression Dalam Pemodelan Indeks Saham Syariah Indonesia Dengan Algoritme Grid Search. *Indonesian Journal of Statistics and Its Applications*, 3(2), 148–160.
- [15] Wandura, W., Cikusin, Y., & Hayat, H. (2021). Wabah Corona Virus (Covid-19) (Studi Pada Desa Pandansari Lor Kecamatan Poncokusumo Kabupaten Malang). *Jurnal Inovasi Penelitian*, 2(5), 1627-1634.
- [16] Yao, Wanxiang & Zhang, Chunxiao & Hao, Haodong & Wang, Xiao & Li, Xianli, 2018. A Support Vector Machine Approach To Estimate Global Solar Radiation With The Influence Of Fog And Haze, *Renewable Energy, Elsevier*, vol. 128(PA), pages 155-162.
- [17] Zhan C, Tse CK, Lai Z, Hao T, Su J 2020. Prediction of COVID-19 spreading profiles in South Korea, Italy and Iran by data-driven coding. *PLoS ONE* 15(7): e0234763