

**PEMODELAN PREDIKTIF PADA KASUS COVID-19 DI  
INDONESIA MENGGUNAKAN *SUPPORT VECTOR  
REGRESSION***



TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar  
Magister Ilmu Komputer (M.Kom)

SAFITRI LINAWATI

14002247

Program Studi Ilmu Komputer (S2)

Sekolah Tinggi Manajemen Informatika Dan Komputer

Nusa Mandiri

2020

## **SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISME**

Yang bertanda tangan di bawah ini :

Nama : Safitri Linawati  
NIM : 14002247  
Program Studi : Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : Data Mining

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Pemodelan Prediktif Pada Kasus Covid-19 di Indonesia Menggunakan *Support Vector Regression*” adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan

Jakarta,  
Yang menyatakan,

Materai Rp 6 000,-

Safitri Linawati

## **SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISME**

Yang bertanda tangan di bawah ini :

Nama : Safitri Linawati  
NIM : 14002247  
Program Studi : Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : Data Mining

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Pemodelan Prediktif Pada Kasus Covid-19 di Indonesia Menggunakan *Support Vector Regression*” adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan

Jakarta,  
Yang menyatakan,

Materai Rp 6 000,-

Safitri Linawati

## HALAMAN PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh:

Nama : Safitri Linawati  
NIM : 14002247  
Program Studi : Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Data Mining*  
Judul Tesis : *Pemodelan Prediktif Pandemi Covid-19 Di Indonesia Menggunakan Support Vector Regression*

Telah dipertahankan pada periode 2020-1 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 12 Agustus 2020

### PEMBIMBING TESIS

Pembimbing I : Dr. Agus Subekti, M.T



### DEWAN PENGUJI

Penguji I : Dr. Hilman Ferdinandus Pardede,  
S.T, M.EICT



Penguji II : Dr. Lindung Parningotan Manik,  
M.T.I



Penguji III /  
Pembimbing I : Dr. Agus Subekti, M.T



	<b>LEMBAR BIMBINGAN TESIS</b>
	<b>STMIK NUSA MANDIRI</b>

NIM : 14002247  
 Nama Lengkap : Safitri Linawati  
 Dosen Pembimbing : Dr. Agus Subekti, MT  
 Judul Tesis : **“Pemodelan Prediktif Pada Kasus Covid-19 di Indonesia Menggunakan Support Vector Regression “**

No	Tanggal Bimbingan	Materi Bimbingan	Paraf Dosen Pembimbing
1.	02 April 2020	Pengajuan Judul + Dataset	
2.	11 April 2020	ACC Dataset	
3.	25 April 2020	Pengajuan BAB I dan Pengajuan BAB II	
4.	09 Mei 2020	ACC Pengajuan BAB I dan Pengajuan BAB II	
5.	18 Mei 2020	Pengajuan BAB III dan Pengajuan BAB IV	
6.	31 Mei 2020	ACC Pengajuan BAB III dan BAB IV dan Pengajuan BAB V	
7.	27 Juli 2020	ACC Pengajuan BAB V	
8.	05 Agustus 2020	ACC Keseluruhan	

Catatan untuk Dosen Pembimbing Bimbingan Tesis

- Dimulai pada tanggal : 02 April 2020
- Diakhiri pada tanggal : 05 Agustus 2020
- Jumlah pertemuan bimbingan : 8 (delapan) Kali

Disetujui Oleh,  
**Dosen Pembimbing**



**Dr. Agus Subekti, MT**

## KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadirat Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan Tesis ini dengan baik. Dimana Tesis ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul Tesis, yang penulis ambil sebagai berikut, **“Pemodelan Prediktif Pada Kasus Covid-19 di Indonesia Menggunakan *Support Vector Regression* ”**.

Tujuan penulisan Tesis ini dibuat sebagai salah satu syarat kelulusan program Sarjana Universitas BSI. Sebagai bahan penulisan diambil berdasarkan hasil penelitian (eksperimen), observasi dan beberapa sumber literatur yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan Tesis ini tidak akan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Dwiza Riana, S.SI, MM, M.Kom selaku Ketua Pasca sarjana Magister Ilmu Komputer STMIK Nusa Mandiri.
2. Dr. Agus Subekti, M.T yang sudah dengan setulus hati membimbing, mengarahkan menyumbangkan ide, waktu, dan tenaganya dalam membimbing penulis untuk menyelesaikan tesis ini.
3. Bapak Ir. Naba Aji Notoseputro selaku Direktur AMIK BSI yang telah mengizinkan penulis untuk menyelesaikan pendidikan S2.
4. Rasiyo dan Ritnah selaku Orang tua tercinta yang telah memberikan dukungan material dan moral kepada penulis.
5. Seluruh staf pengajar (dosen) STMIK Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.

6. Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan tesis ini. Penulis menyadari bahwa penulisan  
Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan dimasa yang akan datang.

Akhir kata semoga Tesis ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 06 Agustus 2020

**Safitri Linawati**

**SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH  
UNTUK KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini, saya :

Nama : Safitri Linawati  
NIM : 14002247  
Program Studi : Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : Data Mining  
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*)** atas karya ilmiah kami yang berjudul : “Pemodelan Prediktif Pada Kasus Covid-19 di Indonesia Menggunakan *Support Vector Regression* ” beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau *bentuk-kan*, mengelolanya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis atau pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 06 Agustus 2020

Yang menyatakan,

Safitri Linawati

## ABSTRAK

Nama : Safitri Linawati  
NIM : 14002247  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Data Mining*  
Judul : “Pemodelan Prediktif Pandemi Covid19 Menggunakan *Support Vector Regression*”

Pada tanggal 11 maret 2020 organisasi kesehatan dunia (WHO) menyatakan bahwa wabah covid19 global adalah pandemi karena kecepatan dan skala penularan virus sangat cepat. Kasus pertama covid-19 dilaporkan pada desember 2019 yang berasal dari wuhan cina dan dengan penyebarannya yang cepat covid-19 kini telah menjadi pandemi yang mempengaruhi 972.303 orang dilebih dari 170 negara didunia salah satunya diindonesia. Pandemi covid-19 mulai membuat resah warga indonesia pada bulan maret 2020. Berdasarkan data yang diambil dari tanggl 2 maret hingga 21 juli 2020 telah diprediksi menggunakan metode *SMOreg* dengan model Polykernel. Pada model ini dapat dihasilkan *Root Mean Squared Error* (RMSE) yang lebih rendah pada data sembuh 0.0545 dan *Mean Squared Error* (MSE) 0.003 . Data kasus baru memiliki *Root Mean Squared Error* (RMSE) 0.1076 dan *Mean Squared Error* (MSE) 0.00845 dan untuk data kematian memiliki *Root Mean Squared Error* (RMSE) 0.0845 dan *Mean Squared Error* (MSE) 0.0071.

Kata kunci: Covid19, *SMOreg*, polykernel, RMSE, dan MSE

## ***ABSTRACT***

Name : Safitri Linawati  
NIM : 14002247  
Study program : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Concentration : *Data Mining*  
Title : “*Predictive Modeling Of the Covid19 Pandemic Using Support Vector Regression*”

On March 11, 2020, the World Health Organization (WHO) stated that the global Covid19 outbreak was a pandemic because the speed and scale of virus transmission was very fast. The first case of Covid-19 was reported in December 2019 which came from Wuhan China and with its rapid spread, Covid-19 has now become a pandemic affecting 972,303 people in more than 170 countries in the world, one of which is Indonesia. The Covid-19 pandemic began to worry Indonesian citizens in March 2020. Based on data taken from March 2 to July 21, 2020 it was predicted that using the SMOreg method with the Polykernel model. In this model, a lower Root Mean Squared Error (RMSE) can be obtained for the recovery data of 0.0545 and the Mean Squared Error (MSE) 0.003. New case data has Root Mean Squared Error (RMSE) 0.1076 and Mean Squared Error (MSE) 0.00845 and for mortality data it has Root Mean Squared Error (RMSE) 0.0845 and Mean Squared Error (MSE) 0.0071.

*Keywords: Covid19, SMOreg, polykernel, RMSE, dan MSE*

## DAFTAR ISI

	Halaman
HALAMAN SAMPUL .....	i
HALAMAN JUDUL.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	vi
ABSTRAK.....	viii
ABSTRCT.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xv
<b>BAB 1. PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang Penulisan.....	1
1.2. Identifikasi Masalah.....	2
1.3. Tujuan Penelitian .....	3
1.4. Ruang Lingkup Penelitian.....	3
1.5. Sistematika Penulisan .....	3
<b>BAB II. LANDASAN TEORI .....</b>	<b>5</b>
2.1. Tinjauan Pustaka.....	5
2.1.1. Pengertian Covid.....	5
2.1.2. Pengertian Time Series Forecasting.....	5
2.1.3. Pengertian Prediksi .....	6
2.1.4. Pengertian Praprocessing Data.....	7
2.1.5. Pengertian Normalisasi .....	8
2.1.6. Pengertian <i>SMOreg</i> .....	8
2.1.7. Pengertian <i>Root Mean Squared Error (RMSE)</i> .....	9
2.1.8. Pengertian <i>Mean Squared Error (MSE)</i> .....	9
2.1.9. Linier Regression .....	10
2.1.10. <i>Multilayer Perceptron</i> .....	11
2.2. Tinjauan Studi .....	14
2.2. Tinjauan Objek Penelitian.....	16
<b>BAB III. METODOLOGI PENELITIAN.....</b>	<b>18</b>
3.1. Tahapan Penelitian.....	18
3.1.1. Penjelasan Tahapan Penelitian.....	19

3.2. <i>Support Vector Regression</i> .....	23
<b>BAB IV. HASIL DAN PEMBAHASAN</b> .....	<b>26</b>
4.1. Metode Yang Diusulkan .....	26
4.2. <i>Propocosing</i> Data .....	26
4.3. Normalisasi .....	29
4.4. Penerapan Metode <i>SMOReg</i> Pada <i>WEKA</i> .....	31
4.5. Hasil Prediksi .....	38
4.5.1. Data Sembuh .....	38
4.5.2. Data Kasus Baru .....	39
4.5.3. Data Kematian .....	40
4.6. Grafik Hasil Prediksi .....	41
4.6.1. Grafik Data Sembuh.....	41
4.6.2. Grafik Data Kasus Baru.....	41
4.6.3. Grafik Data Kematian.....	42
4.7. Perbandingan penelitian terkait .....	43
<b>BAB V. PENUTUP</b> .....	<b>45</b>
5.1. Kesimpulan .....	45
5.2. Saran .....	45

**DAFTAR PUSTAKA**

**DAFTAR RIWAYAT HIDUP**

## DAFTAR TABEL

	Halaman
Tabel II.1 Tabel Paper.....	14
Tabel III.1. Sampel Dataset.....	19
Tabel IV.1.Data Sebelum di Normalize.....	29
Tabel IV.2. Data Sesudah di Normalize.....	30
Tabel IV.3. Hasil pengujian smoreg + poly kernel.....	38
Table IV.4. Hasil pengujian smoreg + RBF Kernel.....	38
Tabel IV.5. Hasil pengujian smoreg + PUK.....	38
Table IV.6. Data sembuh perbulan.....	39
Tabel IV.7. Hasil Data Kasus Baru.....	40
Tabel IV.8. Tabel Hasil Data Kematian.....	41
Tabel IV.9. Perbandingan hasil penelitian.....	43

## DAFTAR GAMBAR

	Halaman
Gambar IV.1. Tampilan Awal Aplikasi Weka.....	27
Gambar IV.2. Tampilan Halaman Antarmuka Weka <i>Explorer</i> .....	27
Gambar IV.3. Tampilan Open File .....	28
Gamabr IV.4. Tampilan Dataset Covid-19 .....	29
Gambar IV.5. Tampilan Proses Normalisasi.....	29
Gambar IV.6. Grafik sebelum dinormalize.....	30
Gamabr IV.7. Grafik sesudah dinormalize .....	30
Gamabr IV.8. Tampilan awal time series forecasting.....	32
Gamabr IV.9. Tampilan Kernel .....	32
Gamabr IV.10. Hasil Pengujian SMOReg Data Sembuh .....	34
Gamabr IV.11. Hasil Pengujian SMOReg Kasus Baru .....	35
Gamabr IV.12. Hasil Pengujian SMOReg Kematian .....	37
Gamabr IV.13. Tampilan Grafik data sembuh.....	41
Gamabr IV.14. Tampilan Grafik Kasus Baru .....	41
Gamabr IV.15. Tampilan Grafik Data Kematian.....	42

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Dalam beberapa bulan ini Indonesia sedang diresahkan dengan pandemi COVID-19. Pandemi adalah penyebaran penyakit diseluruh benua dan diseluruh negara di dunia[1]. Pandemi itu sendiri membuat resah seluruh dunia dikarenakan sangat mudah untuk ditularkan, pandemi yang saat ini meresahkan yaitu pandemi COVID-19. Pada tanggal 11 Maret 2020 Organisasi Kesehatan Dunia (WHO) menyatakan bahwa wabah COVID-19 global adalah pandemi karena kecepatan dan skala penularan virus sangat cepat[2]. Pandemi Virus Corona Baru, COVID-19 adalah topik yang paling banyak dibicarakan, baik oleh citra dan media populer, maupun oleh dunia ilmiah. Situasi ini disebabkan ketakutan, ketidakpastian, dan terutama kurangnya pengetahuan tentang penyakit tersebut[3].

Pada 9 April 2020, pandemi penyakit coronavirus 2019 (COVID-19) telah mengakibatkan 1.521.252 kasus dan 92.798 kematian di seluruh dunia, termasuk 459.165 kasus dan 16.570 kematian di Amerika Serikat (1,2). Personel perawatan kesehatan (HCP) adalah pekerja penting yang didefinisikan sebagai orang yang dibayar dan tidak dibayar yang melayani dalam pengaturan perawatan kesehatan yang memiliki potensi untuk terpapar langsung atau tidak langsung ke pasien atau bahan menular (3). Selama 12 Februari – 9 April, di antara 315.531 kasus COVID-19 yang dilaporkan ke CDC menggunakan formulir standar, 49.370 (16%) termasuk data apakah pasien adalah pekerja perawatan kesehatan di Amerika Serikat; termasuk 9.282 (19%) yang diidentifikasi sebagai HCP. Di antara pasien HCP dengan data yang tersedia, usia rata-rata adalah 42 tahun (kisaran interkuartil [IQR] = 32-54 tahun), 6.603 (73%) adalah perempuan, dan 1.779 (38%) melaporkan setidaknya satu kondisi kesehatan yang mendasarinya. Di antara pasien HCP dengan data tentang perawatan kesehatan, rumah tangga, dan paparan komunitas, 780 (55%) melaporkan kontak dengan pasien COVID-19 hanya dalam pengaturan perawatan kesehatan. Meskipun 4.336 (92%) pasien HCP melaporkan memiliki setidaknya satu gejala di antara demam, batuk, atau sesak napas, 8% sisanya tidak melaporkan gejala ini. Sebagian besar HCP dengan COVID-19

(6.760, 90%) tidak dirawat di rumah sakit; namun, hasil yang parah, termasuk 27 kematian, terjadi di semua kelompok umur; kematian paling sering terjadi pada HCP berusia  $\geq 65$  tahun. Temuan awal ini menyoroti bahwa baik Profesi Kesehatan terkena infeksi di tempat kerja atau di komunitas, kesehatan dan keselamatan tenaga kerja nasional yang esensial ini perlu dilindungi [4].

Kasus pertama covid-19 dilaporkan pada desember 2019 yang berasal dari wuhan cina dan dengan penyebarannya yang cepat covid-19 kini telah menjadi pandemi yang mempengaruhi 972.303 orang lebih dari 170 negara didunia salah satunya diindonesia [5]. Pandemi covid-19 mulai membuat resah warga indonesia pada bulan maret 2020.

Covid-19 merupakan jenis virus yang baru sehingga banyak pihak yang tidak tahu dan tidak mengerti cara penanggulangan virus tersebut. Seiring mewabahnya virus Corona atau Covid-19 ke ratusan negara, Pemerintah Republik Indonesia menerbitkan protokol kesehatan. Protokol tersebut akan dilaksanakan di seluruh Indonesia oleh pemerintah dengan dipandu secara terpusat oleh Kementerian Kesehatan RI (2020) [6].

Dengan melakukan penelitian ini diharapkan dapat memperoleh hasil prediksi kapan berakhirnya kasus covid19 kasus covid 19 dengan menggunakan data dari tanggal 2 maret – 21 Juli 2020 yang diolah menggunakan metode *Time Series Forecasting*. Dengan demikian dibuatlah tesis mengenai “PEMODELAN PREDIKTIF PADA KASUS COVID-19 DI INDONESIA MENGGUNAKAN *SUPPORT VECTOR REGRESSION* “

## 1.2. Identifikasi Masalah

Berdasarkan dengan latar belakang diatas, dapat diidentifikasi beberapa masalah dalam memprediksi kasus covid 19 diindonesia antara lain :

- Model Algoritma SMOreg seperti apa yang Memberikan hasil prediksi paling akurat dalam memprediksi data sembuh, kasus baru dan kematian ?
- Bagaimana prediksi tanggal dimana pandemi akan berakhir ?

### **1.3. Tujuan Penelitian**

Tujuan pada penelitian ini untuk membuat model prediktif dengan jumlah sembuh, kasus baru dan kematian menggunakan metode yang sudah diusulkan. Juga memprediksi tanggal dimana pandemi akan berakhir.

### **1.4. Ruang Lingkup Penelitian**

Agar penelitian ini lebih terarah maka penulis memfokuskan pembahasan hanya membahas pemodelan prediktif dengan jumlah sembuh, kasus baru dan kematian pada covid 19 dan juga hanya membahas mengenai prediksi tanggal dimana pandemi akan berakhir .

### **1.5. Sistematika Penulisan**

#### **BAB I PENDAHULUAN**

Pada bab ini membahas mengenai latar belakang, identifikasi masalah, tujuan penelitian ruang lingkup penelitian, hipotesis dan sistematika penulisan yang akan dibahas pada bab-bab selanjutnya. Bab ini mencakup keseluruhan bab yang akan dijabarkan .

#### **BAB II LANDASAN TEORI**

Pada bab ini berisikan tentang semua teori yang akan dibahas dalam penelitian .

#### **BAB III METODE PENELITIAN**

Pada bab ini membahas mengenai metode yang digunakan dan melakukan tahapan-tahapan penelitian.

#### **BA IV HASIL DAN PEMBAHASAN**

Untuk bab ini membahas hasil penelitian-penelitian dan ujicoba yang telah dilakukan sebelumnya.

## BAB V PENUTUP

Pada bab ini berisikan kesimpulan dan saran dari penelitian dan juga berisikan apa saja yang dibutuhkan untuk pengembangan penelitian selanjutnya .

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Tinjauan Studi**

##### **2.1.1. Pengertian covid19**

Coronavirus adalah virus yang relatif besar yang mengandung genom RNA heliks beruntai tunggal dalam amplop membran. Membran virus dipenuhi oleh glikoprotein yang memberikan mahkota seperti penampilan virus corona. Sementara coronavirus menginfeksi keduanya manusia dan hewan [1].

Covid-19 merupakan jenis virus yang baru sehingga banyak pihak yang tidak tahu dan tidak mengerti cara penanggulangan virus tersebut. Seiring mewabahnya virus Corona atau Covid-19 ke ratusan negara, Pemerintah Republik Indonesia menerbitkan protokol kesehatan. Protokol tersebut akan dilaksanakan di seluruh Indonesia oleh pemerintah dengan dipandu secara terpusat oleh Kementerian Kesehatan RI (2020) [2].

Coronavirus adalah sekumpulan virus dari subfamili Orthocoronavirinae dalam keluarga Coronaviridae dan ordo Nidovirales. Kelompok virus ini yang dapat menyebabkan penyakit pada burung dan mamalia, termasuk manusia. Pada manusia, coronavirus menyebabkan infeksi saluran pernapasan yang umumnya ringan, seperti pilek, meskipun beberapa bentuk penyakit seperti; SARS, MERS, dan COVID-19 sifatnya lebih mematikan [3].

Coronavirus atau sering disebut covid19 ini merupakan sebuah virus yang baru yang mengandung genom RNA heliks beruntai tunggal dalam amplop membran. Wabah corona sendiri sudah menyebar ke ratusan negara salah satunya Indonesia. Virus corona menginfeksi pada manusia dan hewan.

##### **2.1.2. Pengertian Time Series Forecasting**

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. 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 [4].

*Forecasting* adalah Pendekatan peningkatan untuk meningkatkan akurasi permukaan yang lebih tinggi telah banyak dieksploitasi dalam beberapa dekade terakhir. Pendekatan peramalan beban memiliki dua kategori yang perama adalah model peramalan tradisional yang sebagian besar didasarkan pada statistik dan yang lainnya didasarkan pada teknologi buatan. Model tradisional dikhususkan untuk mencari hubungan berulang dari beban listrik sendiri diantara periode waktu sebelumnya [5].

*Time Series Forecasting* adalah urutan titik data musiman yang direkam secara berurutan dalam waktu. Selama beberapa dekade terakhir, banyak karya telah dikhususkan untuk mengembangkan dan meningkatkan model peramalan time series musiman [6].

### **2.1.3. Pengertian Prediksi**

Pengertian Prediksi sama dengan ramalan atau perkiraan. Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan

tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan [7] .

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Pengertian prediksi sama dengan perkiraan. Menurut kamus besar Bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu [8].

Prediksi adalah sama dengan ramalan atau perkiraan. Menurut Kamus Besar Bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Kesimpulannya pengertian prediksi secara istilah akan sangat tergantung pada konteks atau permasalahannya. Berbeda dengan pengertian prediksi secara bahasa yang berarti ramalan atau perkiraan yang sudah menjadi pengertian yang baku [9].

#### **2.1.4. Pengertian Praprocessing Data**

Pra-pemrosesan data adalah salah satu langkah terpenting dalam pembelajaran mesin, yang membantu dalam membangun model pembelajaran mesin dengan lebih akurat. Setiap data scientist harus menghabiskan 80% waktu untuk pra-pemrosesan data dan 20% waktu untuk benar-benar melakukan analisis. Pra-pemrosesan data adalah proses pembersihan data mentah. Data dikumpulkan dalam data raw dan dikonversi ke set data bersih. Dengan kata lain, setiap kali data dikumpulkan dari sumber yang berbeda dikumpulkan dalam format raw dan data ini tidak layak untuk analisis. Oleh karena itu, langkah- langkah tertentu dijalankan untuk mengubah data menjadi kumpulan data kecil yang bersih, bagian dari proses ini disebut data pra-pemrosesan [10].

### 2.1.5. Pengertian Normalisasi

Normalisasi adalah proses untuk menyamakan standar semua data yang digunakan. Salah satu cara untuk normalisasi yaitu Min-Max Normalization, yang dapat dijabarkan dengan persamaan [11].

Normalisasi merupakan teknik penskalaan agar data berada pada rentang nilai yang lebih kecil, misalnya [0.0, 1.0] atau [-1, 1]. Normalisasi data mencoba untuk memberikan bobot yang sama terhadap semua atribut. Hal ini bertujuan untuk meningkatkan akurasi dan efisiensi algoritma yang akan digunakan. Normalisasi min-max digunakan pada penelitian ini untuk melakukan transformasi linear pada data asli. Rentang nilai baru yang akan digunakan adalah [0.0, 1.0] [12].

### 2.1.6. Pengertian *SMOreg*

*Support Vector Regression* sebagai perpanjangan dari mesin vektor dukungan dalam regresi telah banyak diterapkan diberbagai bidang. Namun kesalahan tidak dapat dihindarkan karena berbagai alasan seperti perubahan lajes mekanik dalam perilaku sistem yang keliru, pengukuran, dan misporting termotivasi. Untuk mengurangi pengaruh negatifnya beberapa kesalahan dihilangkan secara langsung dan yang lainnya dideteksi oleh strategi itu sendiri [13].

*SMOreg* adalah singkatan dari *Sequential Minimal Optimization* algoritma untuk implementasi *Support Vector Machine* (SVM) untuk regresi. Pada dasarnya digunakan untuk pelatihan SVM. Pelatihan SVM membutuhkan solusi untuk masalah optimisasi pemrograman kuadratik yang sangat besar. Namun, untuk tujuan ini SMO digunakan. SMO pertama membagi masalah QP menjadi serangkaian unit masalah QP kecil. Unit-unit kecil ini selanjutnya diselesaikan secara analitis [14].

*Support Vector Regression* (SVR) adalah pendekatan peramalan nevel dan telah berhasil digunakan untuk memecahkan masalah deret waktu. Namun, aplikasi model SVR dalam prakiraan waktu musiman belum banyak diselidiki. SVR juga digunakan untuk meramalkan data yang sama. [6].

### **2.1.7. Pengertian *Root Mean Squared Error* (RMSE)**

RMSE (Root Mean Square Error) adalah cara umum dipergunakan untuk mengukur kesalahan model dari prediksi data yang bersifat kuantitatif. RMSE dipergunakan untuk mengetahui ukuran sebaran simpangan titik data dari garis regresi linier atau untuk mengetahui konsentrasi data di sekitar garis regresi linier [15].

Root Mean Square Error (RMSE) digunakan untuk mengetahui besarnya nilai penyimpangan antara data prediksi dengan data observasi. RMSE memiliki nilai antara nol sampai tak terhingga, dengan semakin kecil RMSE maka semakin dekat nilai prediksi model terhadap data observasi [16].

Root Mean Square Error (RMSE) adalah perbandingan yang digunakan untuk membandingkan antara nilai yang didapatkan dari hasil analisa menggunakan tujuh metode dan nilai yang sebenarnya diamati dari lokasi studi yang sedang dimodelkan. Standar nilai hasil perhitungan Root Mean Square Error (RSME) yaitu kisaran 0 – 1,0 [17].

### **2.1.8. Pengertian *Mean Squared Error* (MSE)**

*MSE (Mean Square Error)* digunakan untuk mengukur tingkat akurasi berdasarkan quadratic loss. Nilai *MSE* sangat sensitive terhadap nilai kesalahan atau outlier. Jika terdapat salah satu outlier saja nilai *MSE* akan menjadi besar [4].

MSE (Mean Square Error) adalah rata-rata dari kuadrat kesalahan (error) dipergunakan untuk menghitung kesalahan sample data dan tidak dipergunakan untuk estimasi model [15].

Mean square Error (MSE) adalah metode alternative untuk mengevaluasi teknik peramalan masing-masing kesalahan (selisih data actual terhadap data peramalan) yang dikuadratkan, kemudian dijumlahkan dan dibagi dengan jumlah data. Dalam fase forecasting, menggunakan MSE sebagai suatu ukuran ketetapan juga dapat menimbulkan masalah [18].

### 2.1.9 Linier Regression

Linear Regresi merupakan solusi yang cocok untuk digunakan oleh perusahaan multiproduk karena dengan memperkirakan berbagai kombinasi produk, perusahaan dapat memaksimalkan keuntungan serta memperkirakan jumlah produksi yang tepat. Namun dalam menerapkan regresi linear diperlukan pengetahuan mengenai kombinasi produk yang tepat, target pasar, serta jumlah permintaan terhadap produk itu sendiri [19]. Rumus untuk Linear Regression dengan metode kuadrat terkecil adalah :

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \quad (1)$$

a =

$$n(\sum x^2) - (\sum x)^2 \quad (1)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$y = a + b \cdot x_1 + b \cdot x_2 \quad (3)$$

Keterangan

y : Volume Penjualan

x<sub>1</sub> : Biaya Iklan x<sub>2</sub> : Jumlah Sales

a : Konstanta yang menunjukkan besarnya nilai y apabila x = 0

b : besaran perubahan nilai y

Analisis regresi mempelajari bentuk atau pola hubungan dua variabel atau lebih, yang sebuah variabel diantaranya berperan sebagai variabel bebas, sedangkan variabel lainnya merupakan variabel respon. Secara definisi, regresi adalah bagian ilmu statistika yang mempelajari hubungan fungsional antara satu variabel tak bebas dengan satu variabel bebas atau lebih. Sedangkan, korelasi adalah ilmu yang bertujuan untuk mengetahui seberapa kuat hubungan antara variabel bebas dan tidak bebas. Terdapat perbedaan mendasar antara korelasi dan regresi. Analisis korelasi digunakan untuk mencari arah dan kuatnya hubungan antara dua variabel atau lebih, baik hubungan yang bersifat simetris, kausal dan reciprocal, sedangkan analisis regresi digunakan untuk memprediksi seberapa jauh perubahan nilai variabel independen dimanipulasi atau dinaik turunkan [20].

Persamaan umum regresi sederhana adalah :

$$Y = a + b X \quad (i)$$

$$a = \frac{\sum y - b \sum x}{n} \quad (\text{ii})$$

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad (\text{iii})$$

Menghitung nilai korelasi dengan rumus :

$$r = (\text{iv})$$

Koefisien korelasi (r), nilai korelasi dapat berkisar antara +1 dan -1 ( $-1 \leq r \leq +1$ ) artinya :

- A. Jika  $r = +1$ , maka hubungan sempurna secara positif dan jika mendekati +1 hubungannya sangat kuat dan positif.
- B. Jika  $r = -1$ , maka hubungan sempurna secara negatif dan jika mendekati -1 hubungannya sangat kuat dan negatif.
- C. Jika  $r = 0$ , tidak ada hubungan sama sekali antara variabel X dan variabel Y.
- D. Jika  $r = 1$  atau  $r = -1$ , telah terjadi hubungan linier sempurna yaitu berupa garis lurus. Untuk r yang semakin mengarah ke 0, garis semakin tidak lurus.

Manfaat analisa Regresi Linear antara lain :

- A. mengetahui pengaruh suatu atau beberapa variabel predictor terhadap variabel respon
- B. untuk membuat keputusan apakah naik turunnya variabel dependen dapat dilakukan melalui peningkatan variabel independen atau tidak
- C. untuk mengukur kekuatan hubungan antara variabel respon dan variabel predictor.

### 2.1.10 *Multilayer Perceptron*

Multilayer Perceptron adalah contoh dari ANN yang biasanya digunakan untuk memberikan solusi untuk masalah yang berbeda, misalnya untuk pengenalan pola dan interpolasi . Diperlukan beberapa langkah untuk menjalankan klasifikasi ANN menggunakan arsitektur MLP, yaitu dimulai dengan pengumpulan data, kemudian membuat dan mengkonfigurasi jaringan. Selanjutnya menginisialisasi bobot dan bias. Setelah jaringan dapat melakukan pelatihan, validasi data dan digunakan selama klasifikasi. Kelemahan perceptron adalah tempat perceptron hanya dapat memecahkan masalah yang dapat dipisahkan secara linear. Secara

umum ANN adalah sebuah unit proses yang memiliki input dan mengeluarkan output, dengan

neuron-neuron diorganisasikan sebagai layer. Output dari ANN dapat dikomputasi dengan rumus sebagai berikut

$$O = W_{io} \cdot I$$

(1)

di mana  $W_{io}$  merupakan matriks beban (weight matrix) dengan ukuran  $i \times o$ , dengan  $i$  jumlah input node,  $o$  jumlah output node, input vector  $I$ , dan output vector  $O$ . Secara umum data dipresentasikan dalam input layer, kemudian network akan melakukan proses input dengan

melakukan mengalikan input dengan layer weight (beban). Adapun untuk mempermudah dalam memahami cara kerja MLP, dapat digunakan algoritma yang telah disampaikan pada, sebagai berikut[21].

1. Melakukan inisialisasi network, dengan semua beban diset secara random antara angka -1 sampai dengan +1.
2. Mempresentasikan pola pelatihan pertama pada network yang ada, dan menyimpan hasil output.
3. Membandingkan output network tersebut dengan output target yang ada.
4. Memperbaiki eror secara backward

a. Perbaikan pada layer beban dari output:

$$\omega_{ho} = \omega_{ho} + (\eta \delta_o o_h),$$

di mana  $\omega_{ho}$  merupakan nilai beban dari unit  $h$  yang tersembunyi dengan output unit  $o$ ,  $\eta$  merupakan rasio pelatihan, dan  $o_h$  merupakan keluaran dari unit  $h$  yang tersembunyi. Pada persamaan tersebut,

b. Perbaikan pada beban input:

$$\omega_{ih} = \omega_{ih} + (\eta \delta_h o_i),$$

di mana  $\omega_{ih}$  merupakan nilai beban dari unit  $h$  yang tersembunyi dengan input unit  $i$ ,  $\eta$  merupakan rasio pelatihan, dan  $o_i$  merupakan input dari node  $i$ , dengan

$$\delta_h = o_h (1 - o_h) \sum (\delta_o \omega_{ho}).$$

5. Melakukan perhitungan eror, dengan menghitung rata-rata dari nilai target dan output vector. Fungsi berikut dapat digunakan untuk menghitung eror tersebut.

$$E = \frac{\sqrt{\sum_{n=1}^p (t_0 - o_0)^2}}{p},$$

di mana P merupakan jumlah unit pada output layer.

6. Mengulangi langkah 2 untuk setiap pola pada dataset pelatihan untuk melengkapi satu epoch.
7. Melakukan pertukaran dataset pelatihan secara random. Hal ini untuk mengurangi kemungkinan network dipengaruhi oleh urutan pada data.
8. Mengulangi langkah 2 untuk sejumlah epochs atau hingga error mulai berubah.

A perceptron is a fundamental block of MLP architecture, which is inspired by brain neuron cell. A perceptron receives inputs from the previous layer and forwards the output to next layer after performing some mathematical operations. Equation (1) shows the mathematical operation performed by a perceptron.

$$Y_k(x) = f \left\{ \sum_{i=1}^n (w_{ki}x_i + b_k) \right\}$$

where,  $Y_k$  is a output of kth perceptron,  $w_{ki}$  is ith element of the pre-trained weight matrix of kth perceptron in lth layer,  $x_i$  is ith input of perceptron and  $b_k$  is the bias of a perceptron and  $f$  is the activation function. In this work, we are using the sigmoid function as the non-linear activation function and it is the most computationally intensive part in the perceptron design. Therefore, we used the approximation of this function known as the PLAN function ( $f$ ).

A MLP model contains multiple layers of perceptrons connected in a feed-forward manner. The output of the final layer is terminated with a soft-max function to use it as a classifier. As our implementation is limited to the testing phase of MLP, we used normal maximum instead of the soft-max function to reduce the FPGA resource utilization. The proposed MLP design uses three layers i.e. input, output and a hidden layer. As shown in the Figure 2, Layer 1 is the input layer that performs normalization on all  $i$  number of features. Generally, MLP consist of one or more

number of hidden layers, but in the present work, we used only one hidden layer to reduce the classification delay and hardware resource utilization[22].

## 2.2. Tinjauan Studi

Penelitian ini mengacu pada penelitian-penelitian sebelumnya yang memang sudah membahas mengenai covid19, namun masih sedikit penelitian yang membahas mengenai prediksi covid19 yang ada di Indonesia. Beberapa penelitian terdahulu yang membahas mengenai prediksi dan analisis tentang covid19 dapat dilihat pada tabel dibawah ini .

Tabel 2.1. Tabel Paper

No	Paper	Uraian
1	<i>A python based support vector regression model for prediction of COVID19 cases in India.</i> [23]	Pada penelitian ini membahas mengenai prediksi kasus covid19 di negara India. Penelitian ini menggunakan model SVM untuk memprediksi jumlah total kematian, kasus pulih, jumlah kumulatif dan jumlah kasus harian. Penelitian ini hanya memprediksi untuk periode waktu 1 maret hingga 20 april 2020. Dalam parameter yang digunakan penelitian ini mendapatkan akurasi sebesar 97%. Paper ini hanya membahas mengenai tingkat akurasi saja.
2	<i>From SARS to MERS, thrusting Coronavirus into the Spotlight</i>	Dalam penelitian ini mencoba menganalisis virus covid19 pada hewan. Paper tersebut hanya fokus pada pemahaman tentang epidemiologi, patogenesis, pencegahan dan pengobatan SARS-Cov dan MERS-Cov, serta memberikan rincian tentang struktur penting dan fungsi protein lonjakan pada permukaan masing-masing virus. Pada penelitian ini membandingkan patogenesis pada hewan .

3	<i>The Reproductive number of COVID 19 is higher compared to SARS Coronavirus</i>	<p>Penelitian ini bertujuan untuk mengetahui jumlah reproduksi covid 19 dengan coronavirus SARS mana yang lebih tinggi.</p> <p>Dalam penelitian ini menemukan bahwa estimasi rata-rata R0 untuk covid19 adalah sekitar 3,28 dengan median 2,79 dan IQR 1,16 yang jauh lebih tinggi dari perkiraan WHO di 1,95.</p>
4	<i>Analysis and Forecast of covid 19 spreading in china, italy, and france</i>	<p>Pada penelitian ini menganalisis dinamika temporal penyakit koronavirus 2019 di china, italia dan prancis pada periode waktu 22 januari hingga 15 september 2020. Analisis pertama menunjukkan beberapa universalitas pada penyebaran epidemi bahwa dalam bidang rata-rata sederhana dapat berarti. Penelitian ini menganalisis pada data yang sama dalam model kematian yang mudah terinfeksi yang dipulihkan menunjukkan parameter kinetik yang menggambarkan tingkat pemulihan tampaknya sama. Paper ini menempatkan puncak italia sekitar 21 maret 2020 dengan jumlah puncak individu yang terinfeksi sekitar 26.000 (tidak termasuk yang pulih dan mati) sejumlah kematian sekitar 18.000). Sejak kasus yang dikonfirmasi terinfeksi menjadi 10 dan 20 % dari jumlah sebenarnya, tingkat kematian yang jelas dari covid19 turun antara 4%.</p>
5	<i>Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human</i>	<p>Pada penelitian ini terfokus pada strategi antivirus yang melibatkan molekul kecil dan biologi menargetkan interaksi molekuler kompleks yang terlibat dalam infeksi dan replika virus corona. Analisis paten</p>

	<i>Coronavirus Diseases.</i> [1]	coronavirus biologi meliputi antibodi, sitokin, dan terapi berbasis asam nukleat menargetkan eksperimen gen virus serta berbagai jenis vaksin. Lebih dari 500 paten diungkapkan metodologi keempat biologik ini dengan potensi untuk mengobati dan mencegah coronavirus infeksi yang mungkin berlaku untuk covid-19.
6	<i>Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing.</i> [24]	Penelitian ini terfokus pada pertumbuhan epidemi dan merancang strategi kebijakan untuk mengelola penyebaran covid19 dengan metode machine learning dan cloud computing. Paper ini menunjukkan bahwa menggunakan bobot iteratif untuk pemasangan weibull generalized inverse, kecocokan yang lebih dapat diperoleh untuk mengembangkan kerangka prediksi. Penelitian ini menggunakan platform komputasi awan untuk prediksi perilaku pertumbuhan epidemi yang lebih akurat dan real time. Pendekatan yang digunakan oleh paper tersebut berbasis data yang memiliki akurasi yang lebih tinggi seperti disini dapat sangat berguna repons proaktif dari warga dan pemerintah .

### 2.3. Tinjauan Objek Penelitian

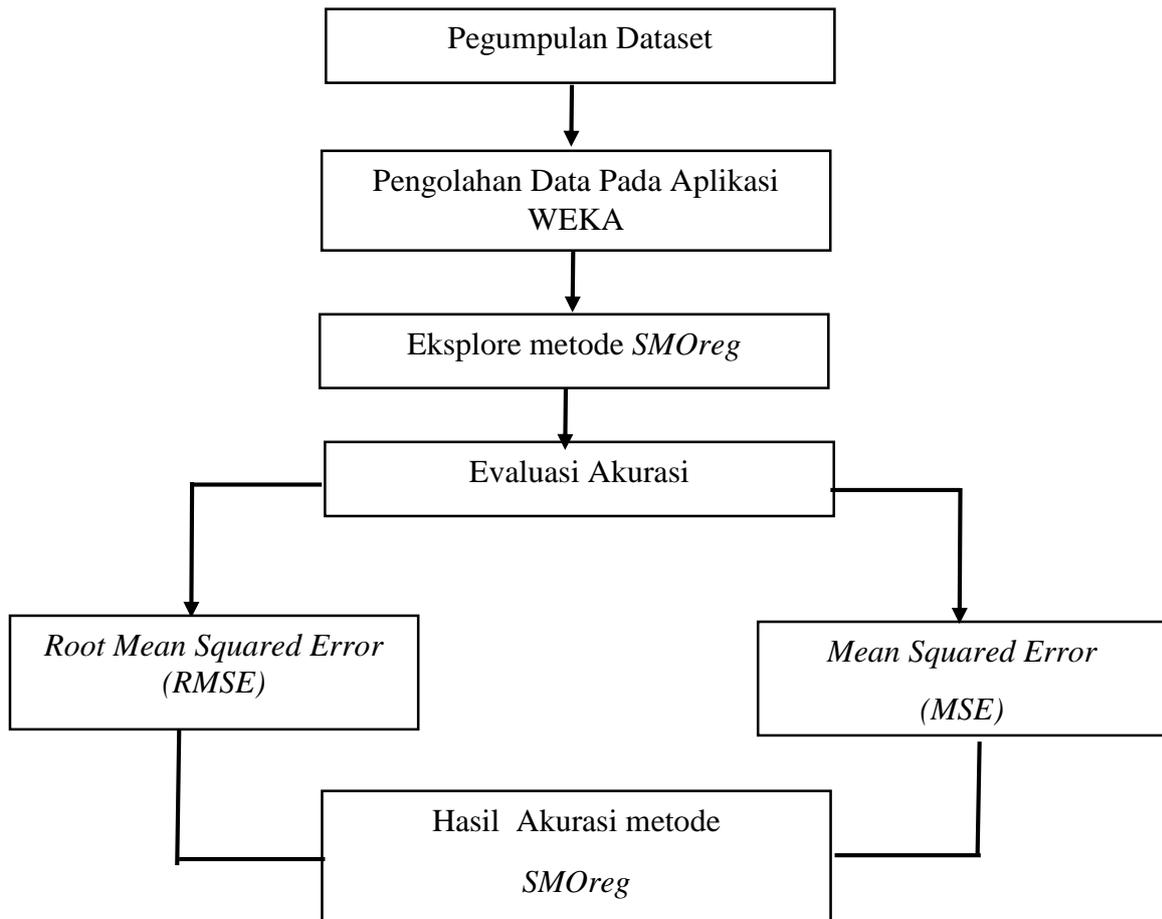
Pada penelitian ini menggunakan dataset publik dari website <https://kawalcovid19.id/> yang memang dibuat untuk update data kasus covid setiap hari . Dataset yang diambil yaitu tanggal, sembuh, kematian dan kasus baru untuk periode 2 Maret – 21 Juli 2020.



**BAB III**  
**METODE PENELITIAN**

**3.1. Tahapan Penelitian**

Ada beberapa tahapan yang dilakukan dalam penelitian ini yaitu pada gambar dibawah ini.



Gambar 3.1. Alur Tahap Penelitian

### 3.1.1. Penjelasan Tahap Penelitian

Pada penelitian ini menggunakan ad 141 data yang diambil dari website resmi kawal covid .

Tabel 3.1.Sampel Dataset

No	Sembuh	kasus baru	kematian	tanggal
1	0	2	0	02/03/2020
2	0	0	0	03/03/2020
3	0	0	0	04/03/2020
4	0	0	0	05/03/2020
5	0	2	0	06/03/2020
6	0	0	0	07/03/2020
7	0	2	0	08/03/2020
8	0	13	0	09/03/2020
9	0	8	0	10/03/2020
10	1	0	2	11/03/2020
11	2	35	0	12/03/2020
12	3	27	1	13/03/2020
13	0	21	0	14/03/2020
14	0	17	0	15/03/2020
15	1	38	2	16/03/2020
16	2	55	12	17/03/2020
17	4	82	6	18/03/2020
18	1	60	7	19/03/2020
19	4	81	6	20/03/2020
20	9	64	10	21/03/2020
21	1	65	1	22/03/2020
22	0	106	6	23/03/2020
23	1	105	3	24/03/2020
24	4	103	20	25/03/2020
25	11	153	9	26/03/2020
26	13	109	15	27/03/2020
27	5	130	12	28/03/2020
28	11	129	8	29/03/2020
29	6	114	14	30/03/2020
30	22	149	21	01/04/2020
31	9	113	13	02/04/2020
32	22	196	11	03/04/2020
33	16	106	10	04/04/2020
34	14	181	7	05/04/2020
35	28	218	11	06/04/2020
36	12	247	12	07/04/2020

37	18	218	19	08/04/2020
38	30	337	40	09/04/2020
39	30	219	26	10/04/2020
40	4	330	21	11/04/2020
41	73	399	46	12/04/2020
42	21	316	26	13/04/2020
43	46	282	60	14/04/2020
44	20	297	10	15/04/2020
45	102	380	27	16/04/2020
46	59	407	24	17/04/2020
47	24	325	15	18/04/2020
48	55	327	47	19/04/2020
49	61	185	8	20/04/2020
50	95	375	26	21/04/2020
51	71	283	20	22/04/2020
52	47	357	11	23/04/2020
53	42	436	42	24/04/2020
54	40	396	31	25/04/2020
55	65	275	23	26/04/2020
56	44	214	22	27/04/2020
57	103	415	8	28/04/2020
58	137	260	11	29/04/2020
59	131	347	8	30/04/2020
60	69	433	8	01/05/2020
61	74	292	31	02/05/2020
62	211	349	14	03/05/2020
63	78	395	19	04/05/2020
64	243	484	8	05/05/2020
65	120	367	23	06/05/2020
66	64	338	35	07/05/2020
67	113	336	13	08/05/2020
68	113	533	16	09/05/2020
69	91	387	14	10/05/2020
70	183	233	18	11/05/2020
71	182	484	16	12/05/2020
72	224	689	21	13/05/2020
73	231	568	15	14/05/2020
74	285	490	33	15/05/2020
75	108	529	13	16/05/2020
76	218	489	59	17/05/2020
77	195	496	43	18/05/2020
78	143	486	30	19/05/2020
79	108	693	21	20/05/2020

80	263	973	36	21/05/2020
81	219	634	48	22/05/2020
82	192	949	25	23/05/2020
83	153	526	21	24/05/2020
84	240	479	19	25/05/2020
85	235	415	27	26/05/2020
86	180	686	55	27/05/2020
87	183	687	23	28/05/2020
88	252	678	24	29/05/2020
89	523	557	53	30/05/2020
90	293	700	40	31/05/2020
91	329	467	28	01/06/2020
92	298	609	22	02/06/2020
93	471	684	35	03/06/2020
94	486	585	23	04/06/2020
95	551	703	49	05/06/2020
96	464	993	31	06/06/2020
97	591	672	50	07/06/2020
98	406	847	32	08/06/2020
99	510	1043	40	09/06/2020
100	715	124	36	10/06/2020
101	507	979	41	11/06/2020
102	577	1111	48	12/06/2020
103	563	1014	43	13/06/2020
104	755	857	43	14/06/2020
105	592	1017	64	15/06/2020
106	580	1106	33	16/06/2020
107	540	1031	45	17/06/2020
108	555	1331	63	18/06/2020
109	551	1041	34	19/06/2020
110	534	1226	56	20/06/2020
111	521	862	36	21/06/2020
112	331	954	35	22/06/2020
113	506	1051	35	23/06/2020
114	417	1113	38	24/06/2020
115	791	1178	47	25/06/2020
116	884	124	63	26/06/2020
117	576	1385	37	27/06/2020
118	1027	1198	34	28/06/2020
119	864	1082	51	29/06/2020
120	1006	1293	71	30/06/2020
121	789	1385	58	01/07/2020
122	1072	1624	53	02/07/2020

123	901	1301	49	03/07/2020
124	651	1447	53	04/07/2020
125	886	1607	82	05/07/2020
126	814	1209	70	06/07/2020
127	866	1268	68	07/07/2020
128	800	1853	50	08/07/2020
129	1066	2657	58	09/07/2020
130	878	1611	52	10/07/2020
131	119	1671	66	11/07/2020
132	919	1681	71	12/07/2020
133	1051	1282	50	13/07/2020
134	947	1591	54	14/07/2020
135	1414	1522	87	15/07/2020
136	1295	1574	76	16/07/2020
137	1489	1462	84	17/07/2020
138	1434	1752	59	18/07/2020
139	2133	1639	127	19/07/2020
140	1576	1693	96	20/07/2020
141	1489	1655	81	21/07/2020

Berdasarkan dari gambar 4.1 penjelasan tahapan-tahapan penelitian yang dilakukan adalah sebagai berikut :

1. Pengumpulan dataset

Dataset yang digunakan dalam penelitian ini adalah penelitian publik yang diambil dari website resmi corona virus yang setiap harinya selalu di perbarui. Dalam dataset tersebut diambil kolom sembuh, kasus baru dan kematian. Dataset yang diambil dari tanggal 2 Maret – 21 juli 2020.

2. Pengolahan dataset pada aplikasi WEKA

Dilakukan pengolahan dataset menggunakan aplikasi weka untuk mencari algoritman mana yang memiliki akurasi paling baik dalam memprediksi sembuh, kasus baru dan kematian yang diambil dari 2 maret 21 Juli 2020.

3. Eksplora metode *SMOreg* pada aplikasi weka

Melakukan Eksperimen menggunakan metode *SMOreg* dengan menggunakan model kernel untuk mencari hasil akurasi paling baik

#### 4. Evaluasi Akurasi

Melakukan evaluasi dari hasil eksperimen yang telah dilakukan dengan ukuran kedekatan hasil pengukuran dengan nilai sebenarnya atau nilai target dengan melihat nilai *Root Mean Squared Error* (RMSE) dan *Mean Squared Error* (MSE).

#### 5. Hasil Prediksi Metode Paling Akurat

Setelah melakukan beberapa proses maka didapatkan nilai akurasi paling akurat yaitu model *Support Vector Regressi* (SMOreg) dengan *Root Mean Squared Error* (RMSE) dan *Mean Squared Error* (MSE) akurasi yang paling terkecil.

### 3.2. Support Vector Regression

Teknik dukungan vektor regresi didasarkan pada prinsip minimalisasi risiko *structured*. Daripada menemukan masalah empiris, SVR bertujuan untuk meminimalkan batas atas kesalahan generalisasi. Fungsi regresi dihasilkan dengan menerapkan serangkaian fungsi linier dimensi tinggi. Oleh karena itu diberikan satu set data  $(x_i, A_i)_{i=1}^N$  (dimana  $x_i$  adalah vektor input;  $A_i$  adalah nilai aktual dan  $N$  adalah jumlah tal dari pola data) Fungsi Regresi dinyatakan sebagai berikut[1]:

$$G = w\phi(x_i) + b \quad (1)$$

Dimana  $(x)$  menunjukkan fitur input dan  $w$  dan  $b$  adalah konfisien. Koefisien (tongkat  $b$ ) dihitung dengan meminimalkan fungsi risiko yang diatur ditampilkan sebagai persamaan . (2)

$$P(G) = C \frac{1}{N} \sum_{i=1}^N L_{\varepsilon}(A_i, G_i) + \frac{1}{2} \|w\|^2 \quad (2)$$

Dimana

$$L_{\varepsilon}(A_i, G_i) = \begin{cases} 0 & \text{if } |A_i - G_i| \leq \varepsilon \\ |A_i - G_i| - \varepsilon & \text{otherwise} \end{cases} \quad (3)$$

Dimana  $C$  and  $\varepsilon$  adalah parameter yang ditentukan pengguna. Parameter  $\varepsilon$  adalah

perbedaan antara nilai aktual dan nilai yang dihitung dari fungsi regresi. Perbedaan ini dapat diperlakukan sebagai tabung disekitar fungsi regresi. Poin di luar tabung dipandang sebagai kesalahan pelatihan. Dalam persamaan (2),  $L\varepsilon (A_i, G_i)$  disebut fungsi kerugian  $\varepsilon$  – tidak sensitif.

Kerugian sama dengan nol jika nilai diperkirakan dalam tabung  $\varepsilon$  -. Selanjutnya item kedua persamaan (2),  $\frac{1}{2}\|w\|^2$  digunakan untuk memperkirakan kerataan suatu fungsi. Dengan demikian  $C$  adalah parameter yang menentukan trade-off antara risiko empiris dan kerataan model. Dua variabel slack positif ( $\xi_i$  and  $\xi_i^*$ ) mewakili jarak dari nilai aktual ke nilai batas yang sesuai dari  $\varepsilon$ -tube, kemudian diperkenalkan. Dua variabel kendur ini sama dengan nol ketika titik data berada dalam tabung  $\varepsilon$ -. Persamaan (2) kemudian dirumuskan kembali kedalam bentuk terbatas berikut :

$$\begin{aligned} \text{Min :} \quad & \frac{1}{2} \|w\|^2 + C \left( \sum_{i=1}^N (\xi_i + \xi_i^*) \right) \\ \text{subjective to} \quad & w\phi(x_i) + b - A_i \leq \varepsilon + \xi_i^*, \quad i = 1, 2, \dots, N \\ & A_i - w\phi(x_i) - b \leq \varepsilon + \xi_i, \quad i = 1, 2, \dots, N \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, N \end{aligned}$$

(4)

Masalah optimasi terbatas ini diselesaikan dengan bentuk langrangia primal berikut :

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \|w\|^2 + C \left( \sum_{i=1}^N (\xi_i + \xi_i^*) \right) \\ & - \sum_{i=1}^N \beta_i [w\phi(x_i) + b - A_i + \varepsilon + \xi_i] \\ & - \sum_{i=1}^N \beta_i^* [A_i - w\phi(x_i) - b + \varepsilon + \xi_i^*] - \sum_{i=1}^N (\alpha_i \xi_i + \alpha_i^* \xi_i^*) \end{aligned}$$

(5)

Eq. 5 diminalkan sehubungan dengan variabel primer  $w, b, c, c^*$ , dan dimaksimalkan sehubungan dengan pengali langrangian non-negatif dan  $\alpha_i, \alpha_i^*$ .

Akhirnya kondisi karush-kuhn-Tucker diterapkan untuk persamaan (4) dan bentuk dual langrangian diberikan oleh persamaan (6).

$$\begin{aligned}
 \text{Max} \quad & \sum_{i=1}^N A_i(\beta_i - \beta_i^*) - \varepsilon \sum_{i=1}^N (\beta_i + \beta_i^*) \\
 & - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\beta_i - \beta_i^*)(\beta_j - \beta_j^*) K(x_i, x_j) \\
 \text{subjective to} \quad & \sum_{i=1}^N (\beta_i - \beta_i^*) = 0 \\
 & 0 \leq \beta_i \leq C, \quad i = 1, 2, \dots, N \\
 & 0 \leq \beta_i^* \leq C, \quad i = 1, 2, \dots, N
 \end{aligned} \tag{6}$$

Pangganda lagrange dalam persamaan (6) memenuhi kesetaraan  $\beta_i * \beta_i^* = 0$ .  
 Pengganda lagrange  $\beta_i$  dan  $\beta_i^*$  ditentukan dengan vektor bobot optimal dari hyperplane regresi diwakili oleh persamaan (7).

$$w^* = \sum_{i=1}^N (\beta_i - \beta_i^*) K(x, x_i) \tag{7}$$

Dengan demikian, fungsi regresi dinyatakan oleh .

$$G(x, \beta, \beta^*) = \sum_{i=1}^N (\beta_i - \beta_i^*) K(x, x_i) + b \tag{8}$$

$K(x_i, x_j)$  adalah fungsi kernel yang nilainya sama dengan produk dalam dari dua vektor  $\chi_i$  dan  $\chi_j$  dalam fitur  $\phi(\chi_i)$  dan  $\phi(\chi_j)$  yang berarti bahwa  $K(x_i, x_j) = \phi(\chi_i) \cdot \phi(\chi_j)$ . Setiap fungsi yang memenuhi kondisi Mercer dapat berfungsi sebagai fungsi kernel

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1. Metode yang diusulkan

Penelitian ini melakukan beberapa tahapan mulai dari pra proses data, *Base Learner*, *Evaluation* dan *Perform Evaluation*.

Sebelum dataset diuji, dilakukan *preprocessing* terlebih dahulu. Melakukan konversi tipe data dari nominal ke numeric, mengubah tipe data tanggal dari nominal ke date melakukan *replace missing values* dan *normalize*.

*Dataset* diuji dengan menggunakan algoritma *SMOreg* dengan memilih *confidence intervals* sebesar 10% dan memilih *Perform Evaluation*, dan juga memilih *periodicity*.

Berdasarkan hasil dari pengujian algoritma *SMOreg* menghasilkan nilai *Root Mean Squared Error (RMSE)* dan *Mean Squared Error (MSE)*.

Langkah selanjutnya yang dilakukan ialah Teknik kernel dengan model *Poly kernel*, *RBF kernel* dan PUK pada algoritma *SMOreg*.

Setelah melakukan eksperimen dengan metode teknik *kernel* terhadap algoritma *SMOreg* maka didapat hasil pengujian dengan *Root Mean Squared Error (RMSE)* dan *Mean Squared Error (MSE)* terkecil .

#### 4.2.Pra-Proses Data

Didalam pemrosesan data, penulis memasukan 4 kolom yang diolah menggunakan aplikasi WEKA yaitu sembuh, kasus baru, kematian dan tanggal yang dijadikan acuan untuk memprediksi beberapa bulan kedepan. Masing-masing atribut memiliki tipe data untuk atribut sembuh, kasus baru dan kematian memiliki tipe data *numeric* sedangkan untuk tanggal memiliki tipe data *date*.

Setelah semua attribut sesuai dengan yang digunakan di *Forecasting* baru lah dataset tersebut diolah. Dataset dibagi menjadi data uji dan data tes yang secara otomatis dipisah menggunakan aplikasi *WEKA*. Data asli diuji terlebih dahulu untuk mencari tingkan errornya kemudian dilakukan normalisasi. Fungsi normalisasi digunakan untuk memasukkan data dan mengubah nilai minimum dan maximum dalam skala 0 dan 1.

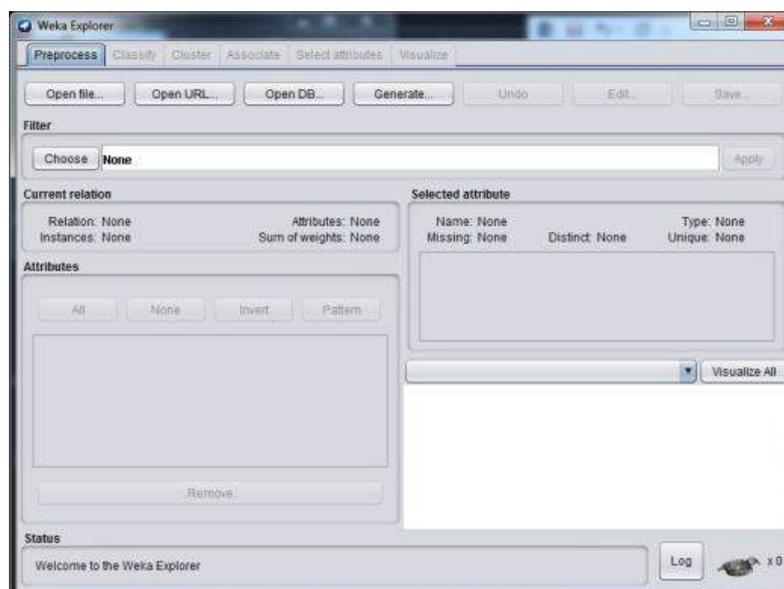
Langkah-langkah Preprocessing data sebagai berikut :

1. Jalankan Aplikasi WEKA



Gambar 4.1. tampilan awal dari aplikasi WEKA .

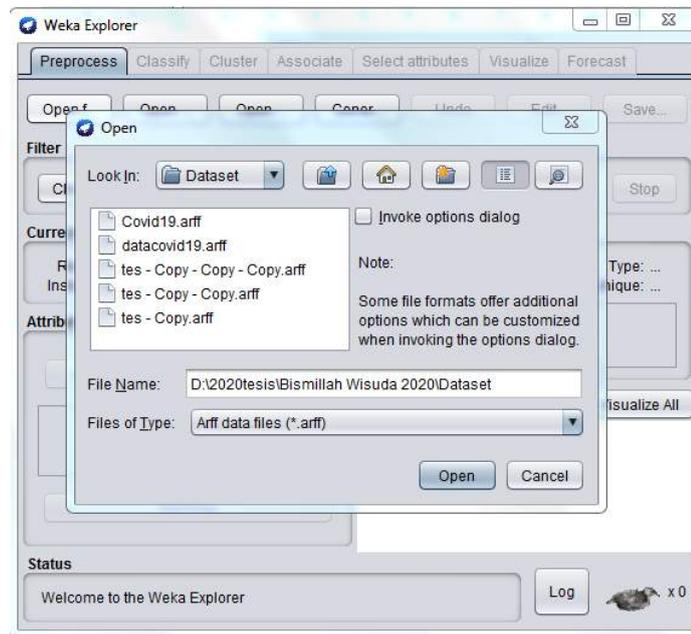
2. Klik tombol Explorer untuk membuka antarmuka *WEKA Explorer*  
 Pada bagian atas terdapat menu Preprocess, Classify, Cluster, Associate, Select Attributes dan Visualize. Namun yang aktif hanya Preprocess, ini dikarenakan sebelum menggunakan algoritma pastikan sudah melakukan set file/memasukan file yang akan diuji.



Gambar 4.2 Tampilan antarmuka aplikasi WEKA

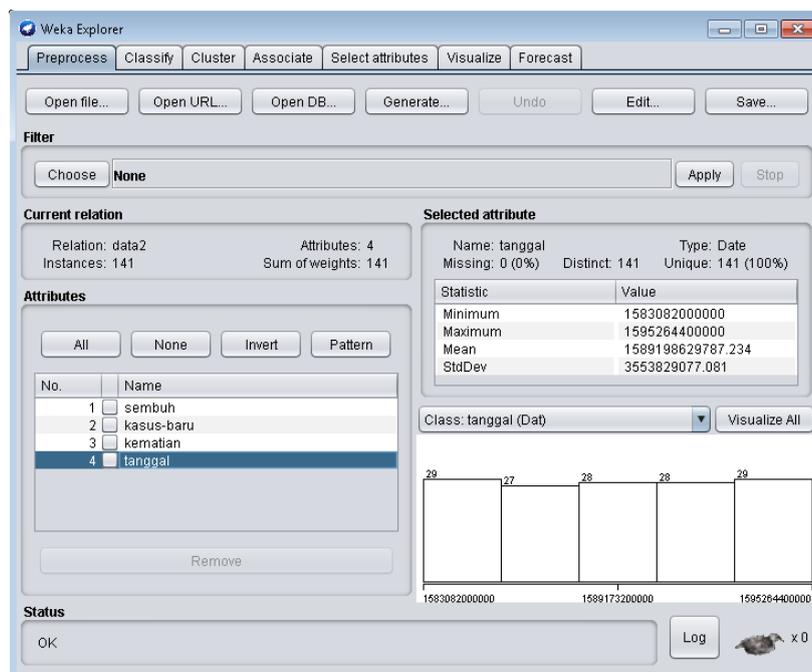
3. Buka File yang akan diuji  
 Pada tahapan ini kita akan membuka file dataset yang akan kita gunakan untuk dilakukan pemrosesan data .

Klik Open file → cari file dataset → Open.



Gambar 4.3. Tampilan Open File

4. Menampilkan atribut yang ada didataset  
 Pada tahapan ini akan menampilkan isi dari dataset, kita bisa mengetahui jumlah *attributes*, *instance*, dan *missing value*. Pada dataset Covid-19 terdapat Relation: data2, Instances/data : 141, Atributtes : 4 yaitu, sembuh, kasus baru, kematian dan tanggal yang memang menjadi patokan untuk memprediksi kasus covid-19.

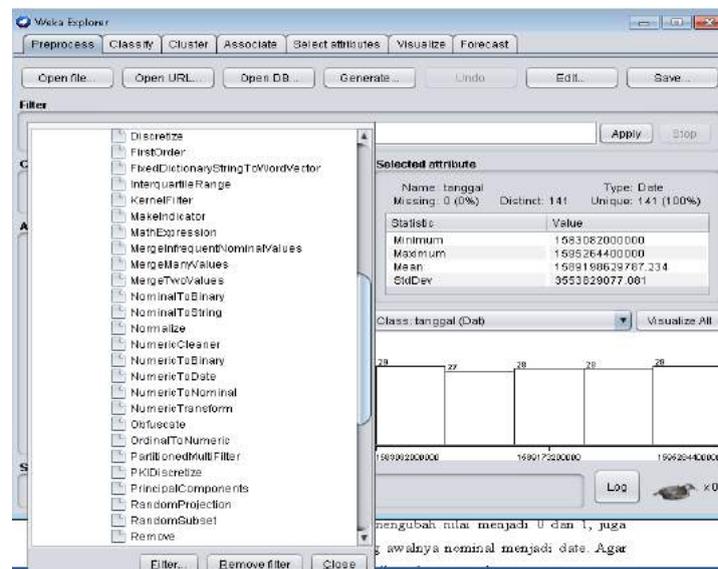


Gambar 4.4. Tampilan Dataset Covid-19

### 4.3. Normalisasi

Data yang dinormalisasi terdapat pada penelitian ini ditampilkan dalam bentuk tabel dan grafik dengan melihat atribut sembuh, kasus baru dan kematian. Data yang ditampilkan data untuk perbulan. Data yang dimasukkan data ssebelum dinormalize dan sesudah dinormalize . Langkah-langkah adalah sebagai berikut

- Klik tombol *Choose* → *Filters* → *Unsupervised* → *Attribute* → *Normalize*



Gambar 4.5. Tampilan Proses Normalisasi

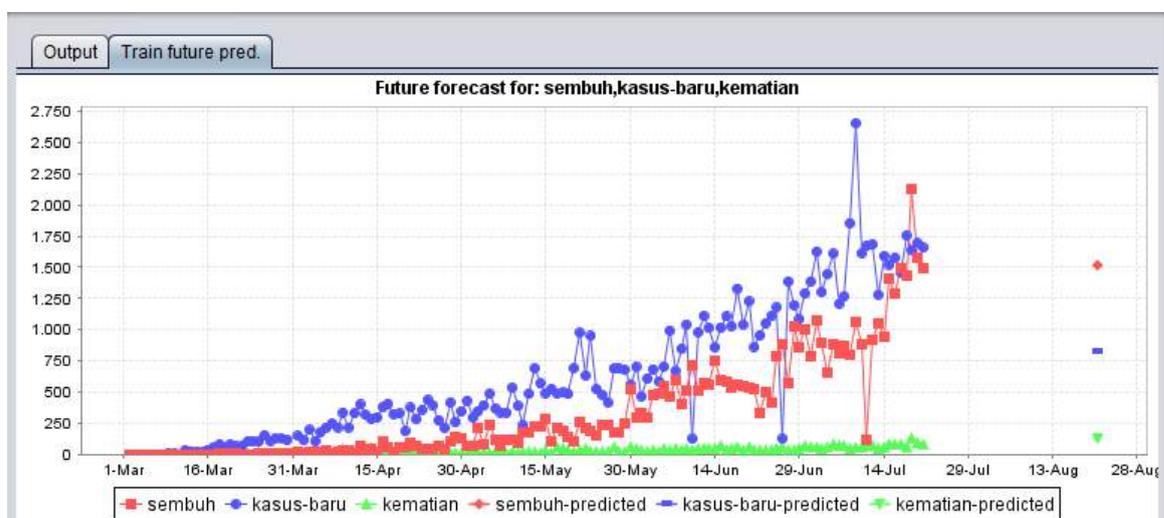
Pada proses normalize ini bertujuan untuk mengubah nilai menjadi 0 dan 1, juga bertujuan untuk mengubah tipe data tanggal yang awalnya nominal menjadi date. Agar bisa diproses kedalam aplikasi weka dan memunculkan data tanggal.

Tabel 4.1 Data sebelum dinormalize

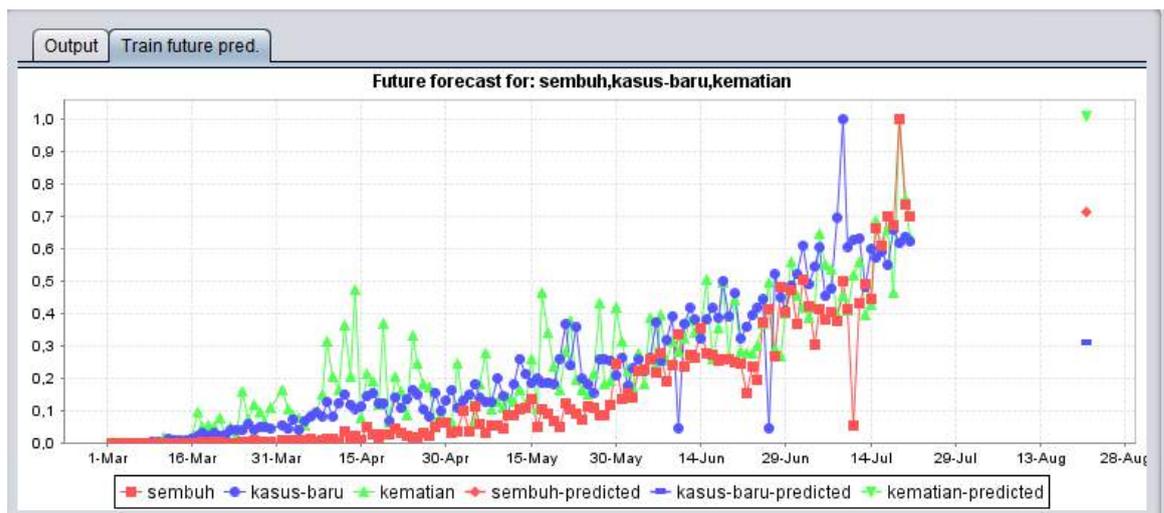
	Linier Regression		SMOreg		Multilayer Perceptron	
	RMSE	MSE	RMSE	MSE	RMSE	MSE
<b>Sembuh</b>	159. 3258	25384. 696	116. 2718	13519. 1271	103. 6774	10748. 9958
<b>Kasus baru</b>	357. 3682	127712. 0219	285. 2617	81374. 2452	403. 5351	162840. 5723
<b>kematian</b>	12. 0554	145. 332	10. 6465	113. 3482	19. 5173	380. 9246

Tabel 4.2 Data sesudah dinormalize

	Linier Regression		SMOreg		Multilayer Perceptron	
	RMSE	MSE	RMSE	MSE	RMSE	MSE
<b>Sembuh</b>	0.0727	0.0053	0.0545	0.003	0.0488	0.0024
<b>Kasus baru</b>	0.1345	0.0181	0.1076	0.0116	0.1519	0.0231
<b>kematian</b>	0.0949	0.009	0.0845	0.0071	0.1539	0.0237



Gambar 4.6. Grafik sebelum dinormalize



Gambar 4.7. Grafik sesudah dinormalize

Berdasarkan dari tabel dan grafik diatas terlihat bahwa metode *SMOreg* memiliki tingkat error terkecil dan mendekati dengan hasil data asli. Sehingga dapat disimpulkan bahwa dalam penulisan ini menggunakan metode *SMOreg*.

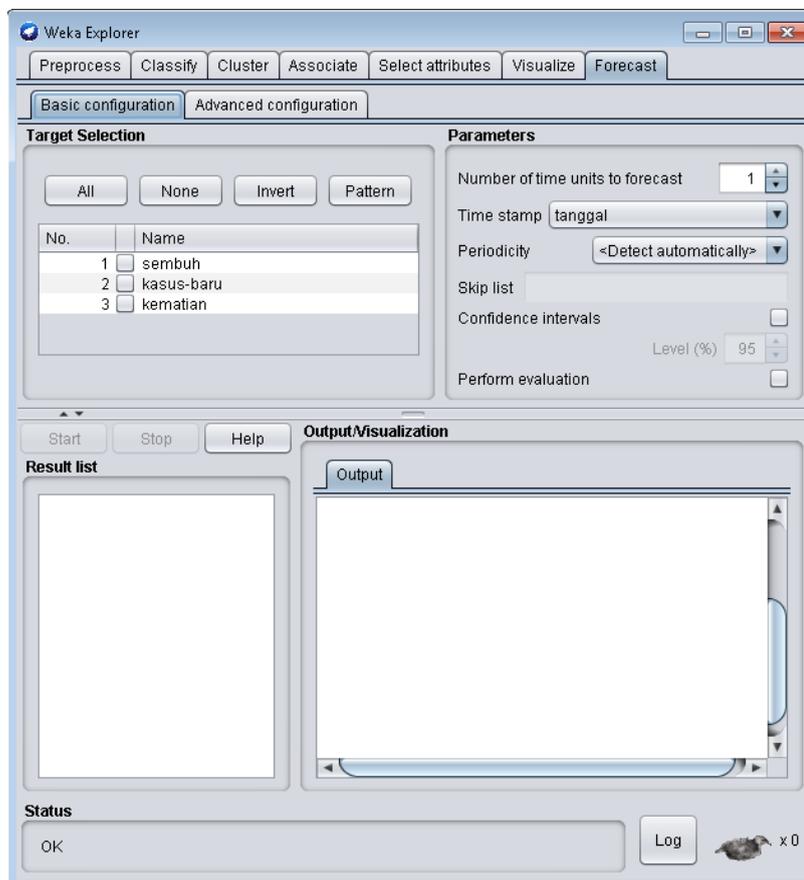
#### 4.4.Penerapan Metode *SMOreg* Pada *WEKA*

Setelah melakukan pra-pemrosan data selanjutnya melakukan pengujian dataset covid19 pada aplikasi weka dengan metode yang telah diinginkan yaitu metode *SMOreg*. Metode *SMOreg* memiliki beberapa pendekatan kernel .

Untuk langkah-langkahnya sebagai berikut :

- *Forecasting* → *Basic Configuration* → *Time stamp* → *periodicity*

Pada gambar dibawah ini ialah tampilan awal dari time series forecasting terdapat 2 menu proses yaitu *Basic Configuration* dan *Advanced Cofiguration*. Dimana pada menu proses *Basic Configuration* berisikan target selection untuk menentukan data mana yang akan diproses lebih dulu dan parameters untuk menentukan berapa waktu target yang dibutuhkan untuk memproses data, time stamp untuk menentukan data mana yang akan digunakan untuk memprediksi, periodicity digunakan untuk menentukan berapa lama data yang akan kita butuhkan .

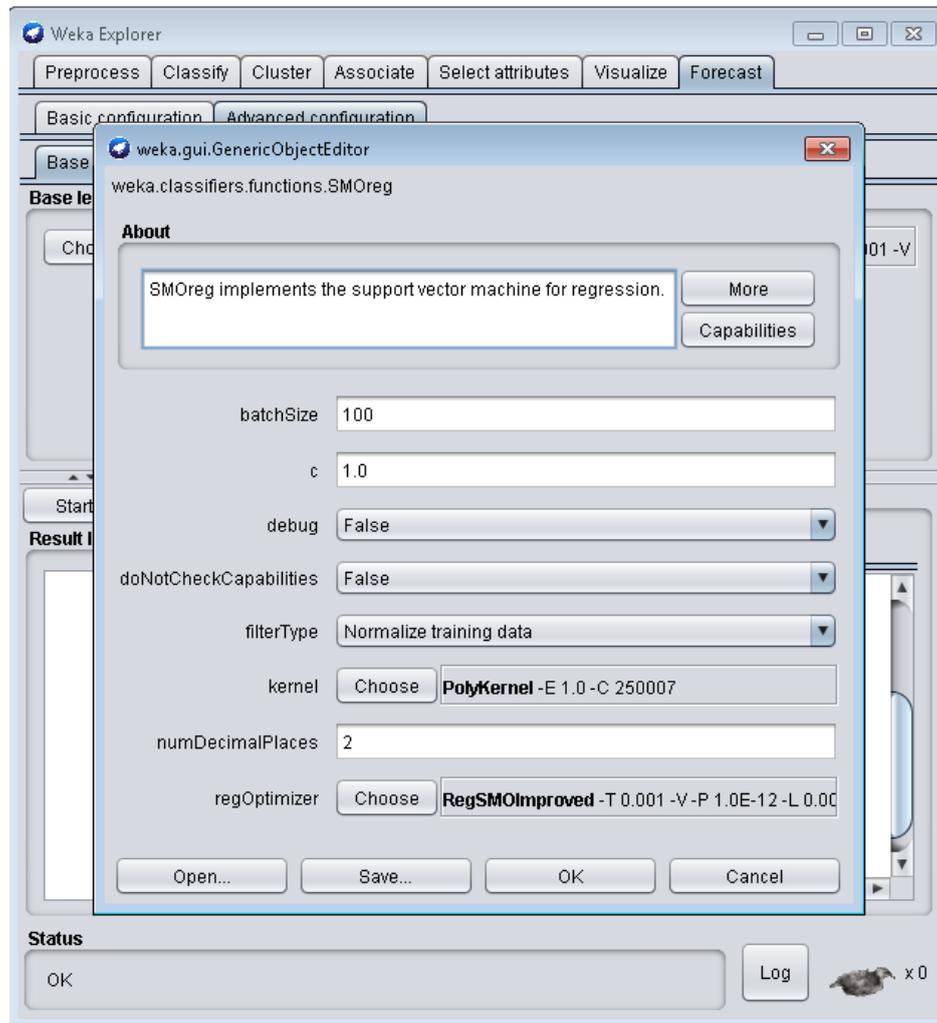


Gambar 4.8. Tampilan awal time series forecasting

Pada penelitian ini *SMOreg* menggunakan beberapa kernel untuk mendapatkan hasil *Root Mean Squared Error (RMSE)* dan *Mean Squared Error (MSE)* terendah. Kernel yang digunakan pada penelitian ini adalah *Poly kernel*, *RBF Kernel*, dan *PUK*.

Langkah-langkah untuk menentukan kernel yaitu sebagai berikut :

- *Advanced Configuration* → *Base Learner* → *Choose* → *SMOReg*



Gambar 4.9. Tampilan Kernel

Pada gambar diatas menentukan kernel mana yang akan digunakan dalam penelitian ini. Penelitian ini menggunakan 3 kernel yaitu *Poly kernel*, *RBF Kernel*, dan *PUK*

Berikut ini adalah pemodelan yang dilakukan menggunakan algoritma *Support Vector Regression* untuk data sembuh, kasus baru dan kematian .

```

sembuh:
SMOreg

weights (not support vectors):
+ 0 * (normalized) Month=jan
+ 0 * (normalized) Month=feb
- 0.0107 * (normalized) Month=mar
- 0.0076 * (normalized) Month=apr
- 0.0274 * (normalized) Month=may
+ 0.0193 * (normalized) Month=jun
+ 0.0264 * (normalized) Month=jul
+ 0 * (normalized) Month=aug
+ 0 * (normalized) Month=sep
+ 0 * (normalized) Month=oct
+ 0 * (normalized) Month=nov
+ 0 * (normalized) Month=dec
- 0.0107 * (normalized) Quarter=Q1
- 0.0157 * (normalized) Quarter=Q2
+ 0.0264 * (normalized) Quarter=Q3
+ 0 * (normalized) Quarter=Q4
+ 0.0253 * (normalized) tanggal-remapped
+ 0.2375 * (normalized) Lag_sembuh-1
+ 0.1039 * (normalized) Lag_sembuh-2
+ 0.0424 * (normalized) Lag_sembuh-3
+ 0.1184 * (normalized) Lag_sembuh-4
- 0.0689 * (normalized) Lag_sembuh-5
- 0.0327 * (normalized) Lag_sembuh-6
- 0.043 * (normalized) Lag_sembuh-7
+ 0.0005 * (normalized) Lag_sembuh-8
+ 0.0499 * (normalized) Lag_sembuh-9
+ 0.0314 * (normalized) Lag_sembuh-10

+ 0.0314 * (normalized) Lag_sembuh-10
- 0.0253 * (normalized) Lag_sembuh-11
+ 0.0554 * (normalized) Lag_sembuh-12
+ 0.1038 * (normalized) Lag_kasus-baru-1
- 0.0065 * (normalized) Lag_kasus-baru-2
- 0.0088 * (normalized) Lag_kasus-baru-3
- 0.0206 * (normalized) Lag_kasus-baru-4
+ 0.055 * (normalized) Lag_kasus-baru-5
+ 0.045 * (normalized) Lag_kasus-baru-6
+ 0.0881 * (normalized) Lag_kasus-baru-7
+ 0.1198 * (normalized) Lag_kasus-baru-8
+ 0.1688 * (normalized) Lag_kasus-baru-9
+ 0.0125 * (normalized) Lag_kasus-baru-10
+ 0.0096 * (normalized) Lag_kasus-baru-11
+ 0.0348 * (normalized) Lag_kasus-baru-12
- 0.0279 * (normalized) Lag_kematian-1
- 0.0059 * (normalized) Lag_kematian-2
+ 0.0059 * (normalized) Lag_kematian-3
+ 0.0377 * (normalized) Lag_kematian-4
+ 0.0407 * (normalized) Lag_kematian-5
- 0.0265 * (normalized) Lag_kematian-6
+ 0.0213 * (normalized) Lag_kematian-7
- 0.0222 * (normalized) Lag_kematian-8
+ 0.0599 * (normalized) Lag_kematian-9
+ 0.0015 * (normalized) Lag_kematian-10
+ 0.0182 * (normalized) Lag_kematian-11
+ 0.0021 * (normalized) Lag_kematian-12
+ 0.0276 * (normalized) tanggal-remapped^2
+ 0.0292 * (normalized) tanggal-remapped^3
+ 0.0859 * (normalized) tanggal-remapped*Lag_sembuh-1
+ 0.0772 * (normalized) tanggal-remapped*Lag_sembuh-2

+ 0.022 * (normalized) Lag_sembuh-10
+ 0.029 * (normalized) Lag_sembuh-11
+ 0.0347 * (normalized) Lag_sembuh-12
+ 0.0282 * (normalized) tanggal-remapped^2
+ 0.0295 * (normalized) tanggal-remapped^3
+ 0.1397 * (normalized) tanggal-remapped*Lag_sembuh-1
+ 0.1871 * (normalized) tanggal-remapped*Lag_sembuh-2
+ 0.0524 * (normalized) tanggal-remapped*Lag_sembuh-3
+ 0.0726 * (normalized) tanggal-remapped*Lag_sembuh-4
- 0.0942 * (normalized) tanggal-remapped*Lag_sembuh-5
- 0.0941 * (normalized) tanggal-remapped*Lag_sembuh-6
- 0.0522 * (normalized) tanggal-remapped*Lag_sembuh-7
- 0.2504 * (normalized) tanggal-remapped*Lag_sembuh-8
- 0.0096 * (normalized) tanggal-remapped*Lag_sembuh-9
+ 0.1035 * (normalized) tanggal-remapped*Lag_sembuh-10
+ 0.0054 * (normalized) tanggal-remapped*Lag_sembuh-11
- 0.0017 * (normalized) tanggal-remapped*Lag_sembuh-12
+ 0.0229

Number of kernel evaluations: 10011 (99.143% cached)

=== Future predictions from end of training data ===
Time      sembuh
2020-03-02      0
2020-04-02      0
2020-05-02      0
2020-06-02      0
2020-07-02      0
2020-08-02      0

```

```

2030-06-02 0.3052
2030-07-02 0.4154
2030-08-02 0.3816
2030-09-02 0.406
2030-10-02 0.3751
2030-11-02 0.4998
2030-12-02 0.4116
2031-01-02 0.0558
2031-02-02 0.4308
2031-03-02 0.4927
2031-04-02 0.444
2031-05-02 0.6629
2031-06-02 0.6071
2031-07-02 0.6981
2031-08-02 0.6723
2031-09-02 1
2031-10-02 0.7389
2031-11-02 0.6981
2031-12-02* 0.6939

=== Evaluation on training data ===
Target          1-step-ahead
=====
sembuh
N                129
Root mean squared error    0.076
Mean squared error        0.0058

Total number of instances: 141

```

Gambar 4.10 Hasil Pengujian Algoritma SMOReg Data Sembuh

Dari gambar diatas dapat dilihat data sembuh dengan menggunakan algoritma *SMOReg* memiliki tingkat error yang rendah

```

kasus-baru:
SMOReg

weights (not support vectors):
+ 0 * (normalized) Month=jan
+ 0 * (normalized) Month=feb
- 0.0322 * (normalized) Month=mar
- 0.0396 * (normalized) Month=apr
- 0.0385 * (normalized) Month=may
+ 0.0542 * (normalized) Month=jun
+ 0.056 * (normalized) Month=jul
+ 0 * (normalized) Month=aug
+ 0 * (normalized) Month=sep
+ 0 * (normalized) Month=oct
+ 0 * (normalized) Month=nov
+ 0 * (normalized) Month=dec
- 0.0322 * (normalized) Quarter=Q1
- 0.0238 * (normalized) Quarter=Q2
+ 0.056 * (normalized) Quarter=Q3
+ 0 * (normalized) Quarter=Q4
+ 0.0676 * (normalized) tanggal-remapped
+ 0.2236 * (normalized) Lag_kasus-baru-1
+ 0.0953 * (normalized) Lag_kasus-baru-2
+ 0.1037 * (normalized) Lag_kasus-baru-3
+ 0.1294 * (normalized) Lag_kasus-baru-4
+ 0.0858 * (normalized) Lag_kasus-baru-5
+ 0.0952 * (normalized) Lag_kasus-baru-6
+ 0.1973 * (normalized) Lag_kasus-baru-7
+ 0.0429 * (normalized) Lag_kasus-baru-8
+ 0.1036 * (normalized) Lag_kasus-baru-9
- 0.0952 * (normalized) Lag_kasus-baru-10

```

```

- 0.0952 * (normalized) Lag_kasus-baru-10
+ 0.0398 * (normalized) Lag_kasus-baru-11
- 0.0512 * (normalized) Lag_kasus-baru-12
+ 0.071 * (normalized) tanggal-remapped^2
+ 0.0726 * (normalized) tanggal-remapped^3
- 0.0747 * (normalized) tanggal-remapped*Lag_kasus-baru-1
- 0.0816 * (normalized) tanggal-remapped*Lag_kasus-baru-2
+ 0.0175 * (normalized) tanggal-remapped*Lag_kasus-baru-3
- 0.2047 * (normalized) tanggal-remapped*Lag_kasus-baru-4
- 0.0182 * (normalized) tanggal-remapped*Lag_kasus-baru-5
- 0.1843 * (normalized) tanggal-remapped*Lag_kasus-baru-6
- 0.0245 * (normalized) tanggal-remapped*Lag_kasus-baru-7
- 0.1733 * (normalized) tanggal-remapped*Lag_kasus-baru-8
+ 0.0404 * (normalized) tanggal-remapped*Lag_kasus-baru-9
+ 0.025 * (normalized) tanggal-remapped*Lag_kasus-baru-10
+ 0.0998 * (normalized) tanggal-remapped*Lag_kasus-baru-11
- 0.0635 * (normalized) tanggal-remapped*Lag_kasus-baru-12
+ 0.0727

```

```

=== Evaluation on training data ===
Target          1-step-ahead
=====
kasus-baru
N                129
Root mean squared error    0.1211
Mean squared error        0.0147

Total number of instances: 141

```

Gambar 4.11. Hasil Pengujian SMOreg data kasus baru

kematian:

SMOreg

weights (not support vectors):

```

+      0      * (normalized) Month=jan
+      0      * (normalized) Month=feb
-     0.0287 * (normalized) Month=mar
-     0.0204 * (normalized) Month=apr
-     0.0291 * (normalized) Month=may
+     0.0375 * (normalized) Month=jun
+     0.0408 * (normalized) Month=jul
+      0      * (normalized) Month=aug
+      0      * (normalized) Month=sep
+      0      * (normalized) Month=oct
+      0      * (normalized) Month=nov
+      0      * (normalized) Month=dec
-     0.0287 * (normalized) Quarter=Q1
-     0.012  * (normalized) Quarter=Q2
+     0.0408 * (normalized) Quarter=Q3
+      0      * (normalized) Quarter=Q4
+     0.0492 * (normalized) tanggal-remapped
+     0.1493 * (normalized) Lag_kematian-1
+     0.2209 * (normalized) Lag_kematian-2
+     0.0893 * (normalized) Lag_kematian-3
+     0.0002 * (normalized) Lag_kematian-4
+     0.1764 * (normalized) Lag_kematian-5
-     0.17   * (normalized) Lag_kematian-6
-     0.0328 * (normalized) Lag_kematian-7
+     0.112  * (normalized) Lag_kematian-8
-     0.0631 * (normalized) Lag_kematian-9
+     0.0922 * (normalized) Lag_kematian-10

```

```

+     0.0922 * (normalized) Lag_kematian-10
+     0.0371 * (normalized) Lag_kematian-11
-     0.0245 * (normalized) Lag_kematian-12
+     0.0513 * (normalized) tanggal-remapped^2
+     0.0522 * (normalized) tanggal-remapped^3
+     0.0789 * (normalized) tanggal-remapped*Lag_kematian-1
-     0.1333 * (normalized) tanggal-remapped*Lag_kematian-2
-     0.2043 * (normalized) tanggal-remapped*Lag_kematian-3
+     0.11   * (normalized) tanggal-remapped*Lag_kematian-4
+     0.0401 * (normalized) tanggal-remapped*Lag_kematian-5
+     0.0952 * (normalized) tanggal-remapped*Lag_kematian-6
+     0.0039 * (normalized) tanggal-remapped*Lag_kematian-7
-     0.1811 * (normalized) tanggal-remapped*Lag_kematian-8
+     0.0186 * (normalized) tanggal-remapped*Lag_kematian-9
-     0.0829 * (normalized) tanggal-remapped*Lag_kematian-10
+     0.0303 * (normalized) tanggal-remapped*Lag_kematian-11
+     0.0004 * (normalized) tanggal-remapped*Lag_kematian-12
+     0.0685

```

```

=== Evaluation on training data ===
Target          1-step-ahead
=====
kematian
  N              129
  Root mean squared error      0.1132
  Mean squared error          0.0128

Total number of instances: 141

```

Gambar 4.12 Hasil Pengujian SMOreg Data Kematian

Tabel 4.3 hasil pengujian smoreg + poly kernel

	Root Mean Squared Error	Mean Squared Error
sembuh	0.0545	0.003
kasus baru	0.1076	0.0116
kematian	0.0845	0.0071

Tabel 4.4 Hasil pengujian smoreg + RBF Kernel

	Root Mean Squared Error	Mean Squared Error
sembuh	0.0797	0.0063
kasus baru	0.0807	0.0065
kematian	0.0966	0.0093

Tabel 4.5 Hasil pengujian smoreg + PUK

	Root Mean Squared Error	Mean Squared Error
sembuh	0.1213	0.0147
kasus baru	0.1131	0.0128
kematian	0.1219	0.0149

Berdasarkan hasil tabel diatas menggunakan beberapa model kernel dapat dilihat bahwa model smoreg dengan menggunakan poly kernel memiliki error terendah. Yaitu untuk data sembuh memiliki *Root Mean Square Error* (RMSE) 0.0545 dan *Mean Squared Error* (MSE) sebesar 0.003, data kasus baru *Root Mean Squared Error* (RMSE) 0.1076 dan *Mean Squared Error* (MSE) sebesar 0.0116, sedangkan untuk data kematian *Root Mean Square Error* (RMSE) 0.0845 dan *Mean Squared Error* (MSE) sebesar 0.0071.

## 4.5 Hasil Prediksi

Untuk melihat prediksi dari masing-masing data dari eksperimen yang telah dilakukan dan dengan model beberapa kernel dan didapatkan model smoreg + polykernel dapat dilihat dalam bentuk tabel dan grafik sebagai berikut .

### 4.5.1. Data Sembuh

**Tabel 4.6 Data sembuh perbulan**

Tanggal	Data actual	Prediksi
02/03/2021	0	156.241
02/04/2021	0	155.538
02/05/2021	1	149.834
02/06/2021	2	150.382
02/07/2021	4	144.331
02/08/2021	1	150.643
02/09/2021	4	157.377
02/10/2021	9	161.332
02/11/2021	1	19.141
02/12/2021	0	175.823
02/01/2022	1	155.738
02/02/2022	4	16.486
02/03/2022	11	140.827
02/04/2022	13	16.915
02/05/2022	5	207.074
02/06/2022	11	203.165
02/07/2022	6	237.118
02/08/2022	22	230.382
02/09/2022	9	589.009
02/10/2022	22	609.522
02/11/2022	16	591.197
02/12/2022	14	65.881
02/01/2023	28	638.203
02/02/2023	12	667.621
02/03/2023	18	669.526
02/02/2029	555	7.061.453

02/03/2029	551	7.464.862
02/04/2029	534	7.128.286
02/05/2029	521	7.013.916
02/06/2029	331	6.502.447
02/07/2029	506	6.730.205
02/08/2029	417	686.715
02/09/2029	791	6.892.213
02/10/2029	884	7.181.392
02/11/2029	576	9.494.334
02/12/2029	1027	9.323.177
02/01/2030	864	10.737.318
02/02/2030	1006	12.254.791
02/03/2030	789	9.534.103
02/04/2030	1072	10.980.587
02/05/2030	901	8.965.852
02/06/2030	651	8.837.311
02/07/2030	886	9.485.809
02/08/2030	814	7.975.547
02/09/2030	866	8.510.822
02/12/2031	1.489	

Berdasarkan data diatas hasil prediksi data sembuh untuk setiap bulannya diambil setiap tanggal 2. Terlihat dari tabel diatas data yang diprediksi sampai 02 desember 2031. Dalam data diatas pada tanggal 02 September 2031 adalah data sembuh yang paling terbanyak.

#### 4.5.2. Data Kasus Baru

**Tabel 4.7 Hasil Data Kasus Baru**

Tanggal	Data Actual	Prediksi
02/03/2021	21	74
02/04/2021	17	1.257.269
02/05/2021	38	1.252.572
02/06/2021	55	130.071
02/07/2021	82	1.333.703
02/08/2021	60	1.470.295
02/09/2021	81	1.432.545
02/10/2021	64	1.541.303

02/09/2029	1178	14.765.453
02/10/2029	124	13.735.555
02/11/2029	1385	1.366.386
02/12/2029	1198	13.825.267
02/01/2030	1082	13.922.075
02/02/2030	1293	14.846.503
02/03/2030	1385	13.882.367
02/04/2030	1624	15.153.277
02/05/2030	1301	12.847.975
02/06/2030	1447	16.407.134
02/07/2030	1607	12.980.815
02/08/2030	1209	14.976.174
02/09/2030	1268	13.070.866
02/10/2030	1853	15.145.413
02/11/2030	2657	14.914.847
02/12/2030	1611	15.596.328
02/01/2031	1671	15.740.273
02/02/2031	1681	15.934.391
02/03/2031	1282	13.767.407
02/04/2031	1591	14.607.867
02/05/2031	1522	1.402.755
02/06/2031	1574	15.494.059
02/07/2031	1462	13.351.242
02/08/2031	1752	1.568.724
02/09/2031	1639	15.558.537
02/10/2031	1693	15.280.034
02/11/2031	1655	14.927.559
2031-12-02	1.413.688	

Dari tabel prediksi diatas dapat disimpulkan bahwa pada setiap bulannya terjadi peningkatan untuk kasus baru. Tetapi ada beberapa data juga yang setiap bulannya terjadi penurunannya untuk lebih detailnya ditampilkan dalam bentuk grafik .

#### 4.5.3. Data Kematian

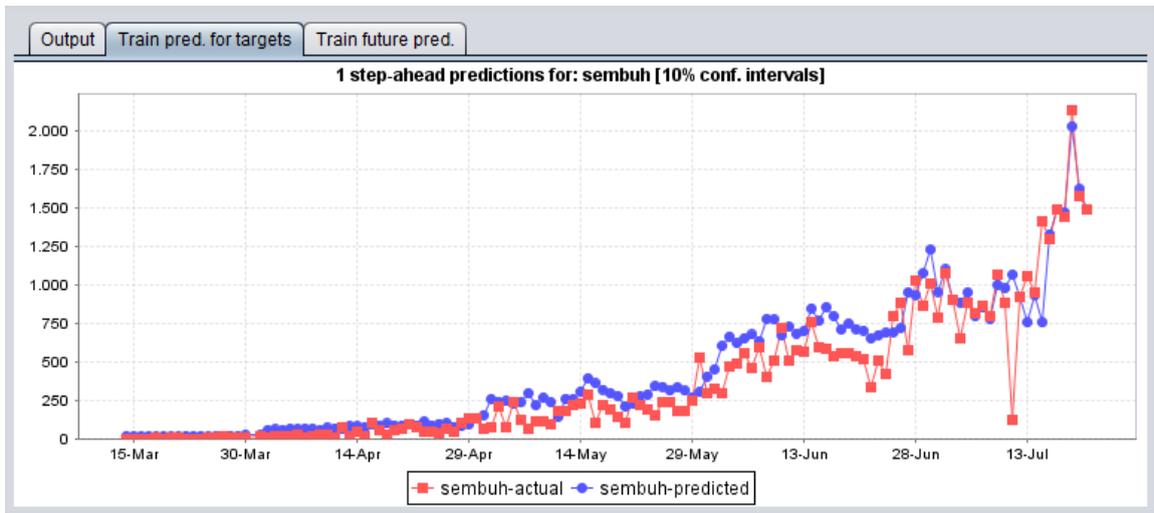
**4.8 Tabel Hasil Data Kematian**

Tanggal	Data Aktual	Prediksi
02/03/2021	0	78.948
02/10/2031	96	831.618
02/11/2031	81	816.547
02/12/2031	615.746	

Dari data diatas juga bisa terlihat untuk data kematian tidak terlalu meningkat juga tidak terllu menurun.

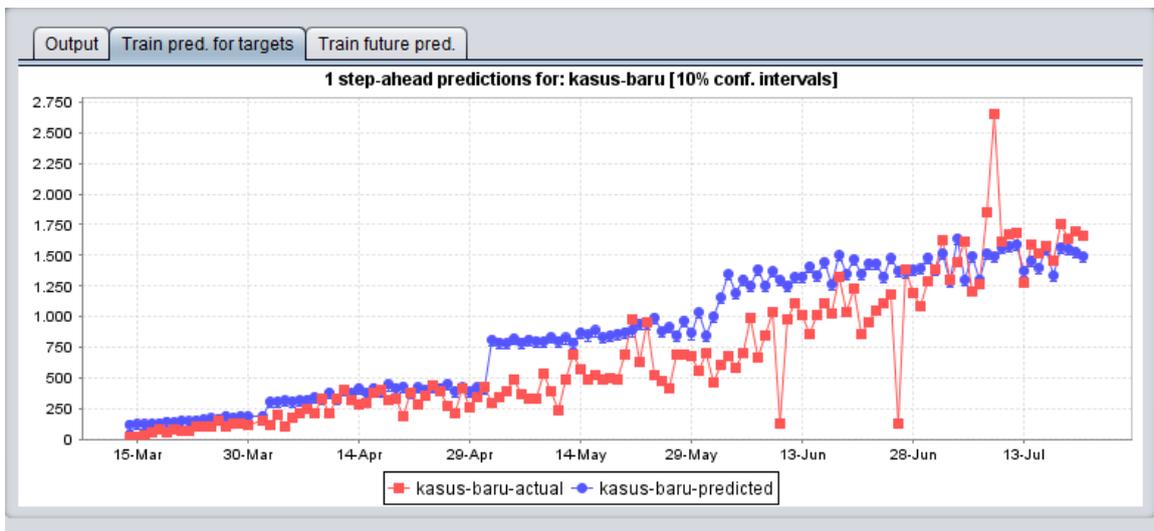
## 4.6.Grafik Hasil Prediksi

### 4.6.1. Grafik Data Sembuh



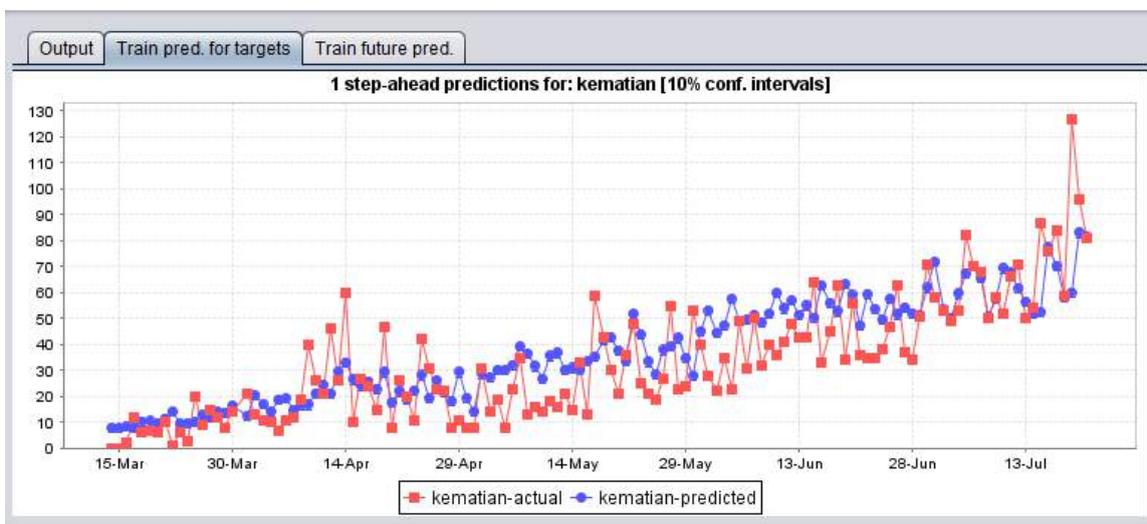
Gambar 4.13 Tampilan Grafik data sembuh

### 4.6.2. Grafik Data Kasus Baru



Gambar 4.14 Tampilan Grafik Kasus Baru

### 4.6.3. Grafik Data Kematian



Gambar 4. 15. Tampilan Grafik Data Kematian

Berdasarkan data grafik diatas dapat dilihat bahwa untuk prediksi sembuh, kasus baru dan kematian tidak jauh berbeda dengan data aslinya. Sehingga dapat diprediksi bahwa setiap data mempunyai prediksi yang tidak signifikan .

### 4.7. Perbandingan penelitian terkait

pada tahap ini menjelaskan tentang penelitian sejenis yang dilakukan oleh penelitian lain.

**Tabel 4.9 Perbandingan hasil penelitian**

	penelitian sebelumnya+ RBF Kernel			Kebaruan Penelitian + PolyKernel		
	RMSE	MSE	Akurasi	RMSE	MSE	Akurasi
<b>sembuh</b>	0.174036	0.030289	97%	0.0545	0.003	99%
<b>kasus baru</b>	0.330830	0.109448	87%	0.1076	0.0845	98%
<b>kematian</b>	0.361727	0.130847	82%	0.0845	0.0071	98%

Perbandingan hasil penelitian yang dilakukan dengan penelitian sebelumnya menunjukkan bahwa kinerja dari model prediksi lebih unggul dengan menggunakan metode *Support Vector Regression* (SMOreg) dengan kernel poly kernel.dan menggunakan normalisasi terlebih dahulu. Pada model ini dapat dihasilkan *Root Mean Squared Error* (RMSE) yang lebih rendah pada data sembuh 0.0545 dan *Mean Squared Error* (MSE) 0.003 dan memiliki akurasi sebesar 99%. Data kasus baru memiliki *Root Mean Squared Error* (RMSE) 0.1076 dan *Mean Squared Error* (MSE) 0.0845 dan memiliki akurasi 98% dan

untuk data kematian memiliki *Root Mean Squared Error* (RMSE) 0.0845 dan *Mean Squared Error* (MSE) 0.0071 memiliki akurasi 98%.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Kesimpulan pada penelitian ini yaitu prediksi jumlah sembuh, kasus baru dan kematian pada covid19 untuk perbulannya. Hasil penelitian yang dilakukan dengan penelitian sebelumnya menunjukkan bahwa kinerja dari model prediksi lebih unggul dengan menggunakan metode *Support Vector Regression* (SMOreg) dengan kernel poly kernel dan menggunakan normalisasi terlebih dahulu. Pada model ini dapat dihasilkan *Root Mean Squared Error* (RMSE) yang lebih rendah pada data sembuh 0.0545 dan *Mean Squared Error* (MSE) 0.003 d. Data kasus baru memiliki *Root Mean Squared Error* (RMSE) 0.1076 dan *Mean Squared Error* (MSE) 0.0845 dan untuk data kematian memiliki *Root Mean Squared Error* (RMSE) 0.0845 dan *Mean Squared Error* (MSE) 0.0071.

Berdasarkan data yang telah diolah dapat di prediksi kapan pandemi akan berakhir yaitu pada tanggal 2 desember 2031 sesuai dengan jumlah yang naik turun tetapi naik turunnya kasus tersebut tidak lah terlalu signifikan.

#### 5.2. Saran

1. Untuk peneliti selanjutnya agar menambahkan beberapa data lagi untuk memprediksi kedepannya .
2. mencoba dengan dataset yang sama tetapi menambahkan data lagi karena data yang digunakan dipenelitian ini hanya sampai tanggal 21 juli 2020 saja.

## DAFTAR REFERENSI

- [1] U. Bhayangkara and J. Raya, "COVID-19 Spread Pattern Using Support Vector Regression," vol. 8, no. 28, pp. 67–74, 2020.
- [2] H. H. Elmousalami and A. E. Hassanien, "Day Level Forecasting for Coronavirus Disease (COVID-19) Spread: Analysis, Modeling and Recommendations," 2020.
- [3] I. J. C. Heart *et al.*, "Coronavirus disease 2019 and cardiovascular system : A narrative review," *IJC Hear. Vasc.*, vol. 29, p. 100557, 2020, doi: 10.1016/j.ijcha.2020.100557.
- [4] C. Liu, Q. Zhou, Y. Li, L. V Garner, S. P. Watkins, L. J. Carter, A. C. Gregg, A. D. Daniels, S. Jervej, and D. Albaiu, "Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases," 2020.
- [5] D. Telaumbanua, "Urgensi Pembentukan Aturan Terkait Pencegahan Covid-19 di Indonesia," vol. 12, no. 1, pp. 59–70, 2020.
- [6] N. R. Yunus, A. Rezki, K. Nabi, M. Saw, M. Wabah, and P. Menular, "No Title."
- [7] B. H. Prakoso, P. Studi, R. Medik, J. Kesehatan, P. N. Jember, and J. Timur, "IMPLEMENTASI SUPPORT VECTOR REGRESSION PADA PREDIKSI," vol. 19, no. 1, pp. 155–162, 2019.
- [8] Z. Zhang and W. Hong, "Electric Load Forecasting by Hybrid Self- Recurrent Support Vector Regression Model With Variational Mode Decomposition and Improved Cuckoo Search Algorithm," vol. 8, 2020.
- [9] P. Pai, K. Lin, C. Lin, and P. Chang, "Expert Systems with Applications Time series forecasting by a seasonal support vector regression model," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4261–4265, 2010.
- [10] P. Pemberian, K. Pinjaman, D. Metode, N. Ajeng, B. W. Sari, M. Kom, D. Prabowo, M. Kom, S. Informasi, U. Amikom, and K. Kunci, "TSUKAMOTO Abstraksi Keywords :," vol. 3, no. 1, 2020.
- [11] S. Kasus, M. Fakultas, and T. Informasi, "PREDIKSI NILAI MAHASISWA MENGGUNAKAN METODE ID3," pp. 144–148, 2019.

- [12] A. Ardiana, “BERDASARKAN KONDISI MUSIM DENGAN PENDEKATAN.”
- [13] R. R. Pratama, “Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia,” vol. 19, no. 2, pp. 302–311, 2020.
- [14] D. A. Mardhika, B. D. Setiawan, and R. C. Wihandika, “Penerapan Algoritma Support Vector Regression Pada Peramalan Hasil Panen Padi Studi Kasus Kabupaten Malang,” vol. 3, no. 10, 2019.
- [15] K. Wisatawan, M. Di, and P. Bali, “Support vector regression,” vol. 11, pp. 60–66, 2017.
- [16] Y. Ye, J. Gao, Y. Shao, C. Li, Y. Jin, and X. Hua, “Robust support vector regression with generic quadratic nonconvex  $\epsilon$ -insensitive loss,” *Appl. Math. Model.*, vol. 82, pp. 235–251, 2020.
- [17] I. Rehman, “Data Mining for Forecasting OGDCL Share Prices Using WEKA,” *2019 13th Int. Conf. Math. Actuar. Sci. Comput. Sci. Stat.*, pp. 1–7, 2019.
- [18] A. Satyo and B. Karno, “Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM ( Long ShortTerm Memory ),” vol. 1, no. 1, pp. 1–8, 2020.
- [19] U. Efendi and A. Kristianto, “UJI PERFORMA ASIMILASI DATA OBSERVASI UDARA ATAS DAN UDARA PERMUKAAN MODEL WRF DI SEMARANG ( STUDI KASUS Sekolah Tinggi Meteorologi Klimatologi dan Geofisika,” 2018.
- [20] E. Suhartanto, L. M. Limantara, H. Arum, and R. Tamaya, “METEOROLOGI KLIMATOLOGI DAN GEOFISIKA SAWAHAN KABUPATEN NGANJUK , JAWA TIMUR,” pp. 1–5, 2011.
- [21] T. P. Kabupaten, “PERBANDINGAN METODE TRIPLE EXPONENTIAL SMOOTHING DAN METODE SEASONAL ARIMA UNTUK PERAMALAN INFLASI DI KOTA Prosiding Sendika : Vol 5 , No 2 , 2019 76 Prosiding Sendika : Vol 5 , No 2 , 2019 77,” vol. 5, no. 2, pp. 76–83, 2019.
- [22] N. Science, C. Phenomena, D. Parbat, and M. Chakraborty, “Chaos , Solitons and Fractals A python based support vector regression model for prediction of COVID19 cases in India,” vol. 138, pp. 3–7, 2020.
- [23] S. Tuli, S. Tuli, R. Tuli, and S. Singh, “Internet of Things Predicting the growth and

trend of COVID-19 pandemic using machine learning and cloud computing R,” vol. 11, 2020.

- [24] P. Pai, K. Lin, C. Lin, and P. Chang, “Expert Systems with Applications Time series forecasting by a seasonal support vector regression model,” *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4261–4265, 2010.

## DAFTAR RIWAYAT HIDUP

### A. Biodata mahasiswa

NIM : 14002247  
Nama Lengkap : Safitri Linawati  
Tempat & Tanggal Lahir : Tebas, 24 Juli 1996  
Alamat Lengkap : Dusun Karya 1, Desa Kuala Dua, Gg Kembang Wonosari  
RT/RW 009/003, Kab Kubu Raya

### B. Pendidikan

1. SDN 20 Kubu Raya, Lulus Tahun 2008
2. SMPN 09 Kubu Raya, Lulus Tahun 2011
3. SMA Kemala Bhayangkari, Lulus Tahun 2014
4. AMIK BSI Pontianak, Lulus Tahun 2017
5. S1 Universitas BSI Bandung, lulus tahun ajaran 2018
6. S2 SMIK Nusa Mandiri Jakarta, Sampai dengan sekarang

### C. Riwayat Pengalaman Berorganisasi / Pekerjaan

1. Asisten Instruktur di Kampus BSI Pontianak 2016/2017
2. Magang dibagian LPPM BSI Pontianak.
3. Magang diBagian Administrasi BSI Dewi Sartika B
4. Instruktur Program Studi Teknologi Komputer Universitas Bina Sarana Informatika 2019 s/d 2020.
5. Dosen Tetap Program Studi Teknik Informatika STMIK Nusa Mandiri 2020 s/d Sekarang.



Jakarta, 06 Agustus 2020

**Safitri Linawati**