



**PREDIKSI KASUS COVID-19 DENGAN METODE
*SUPPORT VECTOR REGRESSION***

(Studi kasus Jawa Timur)

SKRIPSI

Oleh

Muhammad Akbar Firmansyah

171810101038

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER**

2022



**PREDIKSI KASUS COVID-19 DENGAN METODE
*SUPPORT VECTOR REGRESSION***

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat
untuk menyelesaikan studi pada Program Studi Matematika (S-1)
dan mencapai gelar Sarjana Sains.

Oleh
Muhammad Akbar Firmansyah
171810101038

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2022

PERSEMBAHAN

Dengan menyebut nama Allah Yang Maha Pengasih dan Maha Penyayang serta sholawat dan salam selalu tercurahkan kepada Nabi Muhammad SAW, penulis persembahkan skripsi ini sebagai ungkapan kebahagiaan dan rasa terimakasih kepada:

1. Bapak Moch. Duhri, Ibu Sopiya, Adik Mukhamad Iqbal Fadloli serta seluruh keluarga yang selalu mendukung, memberikan doa, kasih sayang, dan motivasi yang tiada henti.
2. Ibu Dian Anggraeni, S.Si., M.Si selaku Dosen Pembimbing Utama dan Bapak Prof. Drs. I Made Tirta, M.Sc, Ph.D selaku Dosen Pembimbing Anggota yang telah membimbing dengan sabar dan telaten dalam penyempurnaan tugas akhir ini.
3. Ibu Dr. Yuliani Setia Dewi, S.Si., M.Si dan Bapak Drs. Moh. Hasan, M.Sc., Ph.D selaku Dosen Penguji yang senantiasa memberikan saran dan kritik demi perbaikan tugas akhir ini.
4. Seluruh jajaran guru MI Walisongo 1 Kedungmaling, SMPN 1 Trowulan, dan MAN 2 Mojokerto yang telah memberikan banyak ilmu kepada penulis.
5. Seluruh teman mulai dari MI hingga sekarang KONIFERTIKA dan semua pihak yang selama ini menjadi bagian hidup penulis yang selalu memberi senyum dan tawa.
6. Himatika “GEOKOMPSTAT”, UKMS TITIK yang telah memberikan banyak pengalaman organisasi dan juga mengajarkan kebaikan.
7. Semua pihak yang membantu penulis dalam menyelesaikan tugas akhir.

MOTTO

Cita-cita itu damai. Sejarah itu penuh kekerasan.

PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Muhammad Akbar Firmansyah

NIM : 171810101038

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “PREDIKSI KASUS COVID-19 DENGAN METODE *SUPPORT VECTOR REGRESSION*” adalah benar-benar hasil karya ilmiah sendiri, kecuali jika dalam pengutipan substansi disebutkan sumbernya dan belum pernah diajukan pada institusi manapun serta bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, Mei 2022

Yang menyatakan,

Muhammad Akbar Firmansyah

NIM 171810101038

SKRIPSI

**PREDIKSI KASUS COVID-19 DENGAN METODE
*SUPPORT VECTOR REGRESSION***

Oleh

Muhammad Akbar Firmansyah

171810101038

Pembimbing

Dosen Pembimbing Utama : Dian Anggraeni, S.Si., M.Si.

Dosen Pembimbing Anggota : Prof. Drs. I Made Tirta, M.Sc, Ph.D.

PENGESAHAN

Skripsi berjudul “PREDIKSI KASUS COVID-19 DENGAN METODE *SUPPORT VECTOR REGRESSION*” telah diuji dan disahkan pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Jember

Tim Penguji :

Ketua,



Dian Anggraeni, S.Si., M.Si.
NIP. 1982021620060420002

Anggota II



Dr. Yuliani Setia Dewi, S.Si., M.Si.
NIP. 197407162000032001

Anggota I,



Prof. Drs. I Made Tirta, M.Sc, Ph.D.
NIP. 195912201985031002

Anggota III



Drs. Moh. Hasan, M.Sc., Ph.D.
NIP. 196404041988021001

Mengesahkan

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Jember

Drs. Achmad Sjaifullah, M.Sc., Ph.D
NIP. 195910091986021001

RINGKASAN

PREDIKSI KASUS COVID-19 DENGAN METODE *SUPPORT VECTOR REGRESSION*; Muhammad Akbar Firmansyah; 171810101038; 48 halaman; Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Covid-19 adalah sebuah penyakit yang disebabkan oleh turunan *coronavirus* baru, ‘CO’ diambil dari *corona*, ‘VI’ *virus*, dan ‘D’ *disease* (penyakit). Sebelumnya, penyakit ini disebut ‘2019 *novel coronavirus*’ atau ‘2019-nCoV’. Covid-19 adalah virus baru yang terikat dengan keluarga virus yang sama yaitu *Severe Acute Respiratory Syndrome* (SARS) dan beberapa jenis virus flu biasa. Covid-19 adalah penyakit menular yang disebabkan oleh sindrom pernapasan akut *coronavirus 2* (Sars-CoV-2). Cepatnya penyebaran penyakit disertai penambahan kasus yang masih terus melonjak, termasuk di Indonesia, serta beragamnya manifestasi klinis, berbagai upaya pencegahan penyebaran virus Covid-19 pun dilakukan oleh pemerintah di negara-negara di dunia guna memutus rantai penyebaran virus Covid-19 ini. Analisis regresi merupakan suatu metode yang digunakan untuk mengetahui hubungan antara suatu variabel dependen terhadap satu atau lebih variabel independen. Analisis regresi dapat membandingkan pengaruh variabel yang diukur pada skala yang berbeda. Metode ini dapat digunakan untuk mengevaluasi kumpulan variabel terbaik yang akan digunakan untuk membangun model prediksi.

Salah satu teknik regresi adalah *Support Vector Regression* (SVR). *Support Vector Regression* merupakan metode pengembangan SVM untuk kasus regresi. Tujuan dari SVR adalah untuk menemukan sebuah fungsi (x) sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah error dan membuat setipis mungkin. Tujuan dari SVR ini adalah untuk memetakan vector input ke dalam dimensi yang lebih tinggi. Penelitian ini menggunakan data harian kasus covid-19 di Jawa Timur sebagai data sekunder. Data ini dapat diakses di <https://www.kaggle.com>. Variabel yang digunakan adalah variabel prediktor dan variabel respon. Variabel respon (y) dalam

penelitian ini adalah total kasus harian di Jawa Timur yang digunakan sebagai target. Variabel prediktor (x) pada penelitian ini sebagai variabel pendukung adalah kasus baru, kematian baru, sembuh baru, kasus aktif baru, total kematian, total sembuh. Dalam pemodelan SVR optimalisasi perlu dilakukan dan validasi model dilakukan dengan kriteria *MAPE* (*Mean Absolute Percentage Error*) yang terkecil. Prediksi dan observasi menunjukkan bahwa hasil prediksi kasus Covid-19 di Jawa Timur mendekati data observasi dengan nilai *MAPE* 6,7%, sehingga model yang diperoleh adalah model yang baik untuk memprediksi kasus Covid-19 di Jawa Timur.

PRAKATA

Puji syukur kehadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul “PREDIKSI KASUS COVID-19 DENGAN METODE *SUPPORT VECTOR REGRESSION*”. Tugas akhir ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan pendidikan strata satu (S-1) pada Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan tugas akhir ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terimakasih kepada:

1. Ibu Dian Anggraeni, S.Si., M.Si selaku Dosen Pembimbing Utama dan Bapak Prof. Drs. I Made Tirta, M.Sc, Ph.D. selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan tugas akhir ini;
2. Ibu Dr. Yuliani Setia Dewi, S.Si., M.Si selaku Dosen Penguji I dan Bapak Drs. Moh. Hasan, M.Sc., Ph.D. selaku Dosen Penguji II yang senantiasa memberikan saran dan kritik dalam penyempurnaan tugas akhir ini;
3. Dosen dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember;
4. Keluarga yang telah memberikan semangat dan doa tulus ikhlas penuh kasih sayang;
5. Teman-teman dan semua pihak yang telah membantu dan memberi semangat.

Semoga bantuan, bimbingan, dan dorongan yang telah diberikan dicatat sebagai amal baik oleh Allah SWT dan mendapat balasan yang sesuai dari-Nya. Penulis juga menerima saran dan kritik dari semua pihak demi kesempurnaan tugas akhir ini. Akhirnya penulis berharap semoga tugas akhir ini bisa bermanfaat.

Jember, Mei 2022

Penulis

DAFTAR ISI

COVER.....	i
PERSEMBAHAN.....	ii
MOTTO.....	iii
PERNYATAAN.....	iv
SKRIPSI.....	v
PENGESAHAN.....	vi
RINGKASAN.....	vii
PRAKATA.....	ix
DAFTAR ISI	x
DAFTAR TABEL....	xii
DAFTAR GAMBAR.....	xiii
BAB 1 PENDAHULUAN....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan.....	3
1.4 Manfaat.....	3
BAB 2 TINJAUAN PUSTAKA.....	4
2.1 Covid-19.....	4
2.2 Analisis Regresi.....	5
2.3 Support Vector Machine.....	5
2.4 Support Vector Regression.....	7
2.5 Pengukuran Performa Hasil Prediksi.....	9
BAB 3 METODE PENELITIAN.....	10
3.1 Data... ..	10
3.2 Langkah Analisis Data.....	10
3.3 Diagram Alir.	11
BAB 4 HASIL DAN PEMBAHASAN.....	12
4.1 Deskripsi Kasus Covid-19 di Jawa Timur	12
4.2 Data Penelitian.....	13
4.3 Prediksi dengan Support Vector Regression (SVR)	13

4.2.1 <i>Tune Parameter SVR</i>	13
4.2.2 Penentuan Fungsi Kernel Sebagai Parameter Terbaik.....	13
4.2.3 Prediksi Kasus Covid-19 dengan Model SVR.....	15
BAB 5 KESIMPULAN	17
5.1 Kesimpulan	17
5.2 Saran	17
DAFTAR PUSTAKA	18
LAMPIRAN	19

DAFTAR TABEL

3.1 Variabel Prediktor	10
4.1 <i>Tune eror</i> dengan 7 parameter <i>cost</i>	13
4.2 Penentuan tipe kernel terbaik data <i>training</i>	14
4.3 Input prediksi SVR.....	16

DAFTAR GAMBAR

2.1 Ilustrasi SVM.....	6
2.2 Ilustrasi SVR.....	7
3.1 Diagram Alir.....	11
4.1 Plot Kasus Covid-19 Jawa Timur	12
4.2 Plot Data Training.....	15
4.3 Plot Data <i>Testing</i>	16

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Kasus pneumonia yang awalnya tidak diketahui penyebabnya pertama kali dilaporkan di Wuhan, Provinsi Hubei, Cina pada Desember 2019. Penyakit ini berkembang sangat pesat dan telah menyebar ke berbagai provinsi lain di Cina, bahkan menyebar hingga ke Thailand dan Korea Selatan dalam kurun waktu kurang dari satu bulan. Pada 11 Februari 2020, *World Health Organization* (WHO) mengumumkan nama penyakit ini sebagai *Virus Corona Disease* (Covid-19) yang disebabkan oleh virus SARS-CoV-2, yang sebelumnya disebut 2019-nCoV, dan dinyatakan sebagai pandemik pada tanggal 12 Maret 2020 (Susilo dkk., 2020). Kasus konfirmasi Covid-19 di Indonesia masih terus bertambah. Berdasarkan laporan Kemenkes RI, pada tanggal 14 Maret 2021 tercatat 1.419.455 juta kasus konfirmasi dan 38.426 ribu kasus meninggal (Kemenkes RI, 2020). Seiring dengan terus meningkatnya kasus terkonfirmasi Covid-19, penelitian mengenai Covid19 masih berlanjut hingga saat ini. Berdasarkan penelitian Xu dkk., (2020) dan Zhu dkk., (2020), ditemukan bahwa agen penyebab Covid-19 berasal dari genus *betacoronavirus*, yang merupakan genus yang sama dengan agen penyebab *Severe Acute Respiratory Syndrome* (SARS) dan *Middle East Respiratory Syndrome* (MERS). Virus dapat melewati membran mukosa, terutama mukosa nasal dan laring, kemudian memasuki paru-paru melalui traktus respiratorius dan selanjutnya menuju organ target (Gennaro dkk., 2020). Covid-19 menjadi perhatian utama dunia karena cepatnya penyebaran penyakit disertai penambahan kasus yang masih terus melonjak, termasuk di Indonesia, serta beragamnya manifestasi klinis. Berbagai upaya pencegahan penyebaran virus Covid-19 pun dilakukan oleh pemerintah di negara-negara di dunia guna memutus rantai penyebaran virus Covid-19 ini yakni *social distancing* dan menjaga protokol kesehatan.

Analisis regresi merupakan suatu metode yang digunakan untuk mengetahui hubungan antara suatu variabel respon terhadap satu atau lebih variabel prediktor (Tukur dan Usman, 2018). Analisis regresi dapat membandingkan pengaruh variabel yang diukur pada skala yang berbeda (Ray, 2015). Metode ini dapat

untuk mengevaluasi kumpulan variabel terbaik yang akan digunakan untuk membangun model prediksi (Ray, 2015).

Salah satu teknik regresi adalah *Support Vector Regression (SVR)*. *Support Vector Regression* merupakan metode pengembangan SVM untuk kasus regresi. Tujuan dari SVR adalah untuk menemukan sebuah fungsi (x) sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah error dan membuat setipis mungkin (Scholkopt dkk, 2002). Abe (2005) menyatakan tujuan dari SVR ini adalah untuk memetakan vector input ke dalam dimensi yang lebih tinggi.

Terdapat sejumlah peneliti yang telah melakukan penelitian mengenai kasus Covid-19. Beberapa penelitian yang telah dilakukan adalah dengan metode SVM yaitu, Hanna (2020) melakukan penelitian menggunakan metode SVM dengan data total kasus Covid-19 di dunia dan di Indonesia. Penelitian ini bertujuan memprediksi banyaknya kasus Covid-19 terutama di Indonesia pada waktu tertentu mendatang. Hasil dari penelitian tersebut adalah tingkat kesalahan yang diperoleh dengan metode SVM tergolong sangat kecil. Peneliti lain yakni Milania (2021) melakukan penelitian menggunakan metode SVR dengan data index harga saham. Penelitian tersebut bertujuan menganalisis prediksi indeks harga saham pada waktu mendatang. Peneliti lain yakni Aldo (2021) melakukan prediksi jumlah kasus Covid-19 menggunakan metode KNN. Hasil dari penelitian tersebut menghasilkan nilai yang sangat baik dengan menambahkan jumlah beberapa data dari jumlah data sebelumnya.

Berdasarkan metode yang telah dijelaskan tersebut, penulis tertarik untuk menganalisis prediksi banyaknya kasus Covid-19 di Jawa Timur menggunakan metode *Support Vector Regression*. Dengan penelitian ini diharapkan pemerintah maupun masyarakat Jawa Timur dan di Indonesia dapat mengetahui perkembangan kasus Covid-19 di wilayahnya sehingga mempermudah pengendalian pandemi ini kedepannya.

1.2 Rumusan Masalah

Bagaimana penerapan metode *Support Vector Regression* dalam memprediksi jumlah kasus Covid-19 di Jawa Timur?

1.3 Tujuan

Berdasarkan rumusan masalah diatas tujuan yang ingin dicapai adalah mendapatkan model untuk mengetahui hasil prediksi jumlah kasus Covid-19 menggunakan metode *Support Vector Regression* di Jawa Timur. Caranya adalah dengan melalui tahap-tahap yang akan dikerjakan.

1.4 Manfaat

Manfaat dari penelitian ini adalah sebagai berikut,

- a. Memberikan informasi kepada masyarakat tentang pertumbuhan kasus Covid-19 di Jawa Timur sehingga masyarakat lebih waspada dan melakukan pencegahan dini.
- b. Mengetahui hasil prediksi kasus Covid-19 menggunakan metode *Support Vector Regression*.

BAB 2. TINJAUAN PUSTAKA

2.1 Covid-19

Covid-19 adalah sebuah penyakit yang disebabkan oleh turunan *coronavirus* baru, ‘CO’ diambil dari *corona*, ‘VI’ *virus*, dan ‘D’ *disease* (penyakit). Sebelumnya, penyakit ini disebut ‘2019 *novel coronavirus*’ atau ‘2019-nCoV’. Covid-19 adalah virus baru yang terikat dengan keluarga virus yang sama yaitu *Severe Acute Respiratory Syndrome* (SARS) dan beberapa jenis virus flu biasa (WHO, 2020). Covid-19 adalah penyakit menular yang disebabkan oleh sindrom pernapasan akut *coronavirus 2* (Sars-CoV-2). Penyakit ini pertama kali ditemukan pada Desember 2019 di Wuhan, ibukota provinsi Hubei China, dan sejak itu menyebar secara global diseluruh dunia dan mengakibatkan pandemi *coronavirus* dari tahun 2019 sampai sekarang. Organisasi Kesehatan Dunia (WHO) mendeklarasikan wabah *coronavirus* sebagai Kesehatan Masyarakat Darurat Internasional (PHEIC) pada 30 Januari 2020, dan menjadi pandemi pada 11 Maret 2020. Berdasarkan penelitian Xu dkk., (2020) dan Zhu dkk., (2020), ditemukan bahwa agen penyebab Covid-19 berasal dari genus *betacoronavirus*, yang merupakan genus yang sama dengan agen penyebab *Severe Acute Respiratory Syndrome* (SARS) dan *Middle East Respiratory Syndrome* (MERS). Virus dapat melewati membran mukosa, terutama mukosa nasal dan laring, kemudian memasuki paru-paru melalui traktus respiratorius dan selanjutnya menuju organ target (Gennaro dkk., 2020). Saat ini Covid-19 menjadi perhatian utama dunia. Cepatnya penyebaran penyakit disertai penambahan kasus yang masih terus melonjak, termasuk di Indonesia, serta beragamnya manifestasi klinis.

Sejak 31 Desember 2019 hingga 3 Januari 2020 kasus ini meningkat pesat, ditandai dengan dilaporkannya sebanyak 44 kasus. Tidak sampai satu bulan, penyakit ini telah menyebar di berbagai provinsi lain di China, Thailand, Jepang, dan Korea Selatan. Virus ini dapat ditularkan dari manusia ke manusia dan telah menyebar secara luas. Oleh karena itu, dibutuhkan kemampuan untuk memprediksi kasus Covid-19 dengan tepat agar masyarakat dapat lebih waspada dengan tidak

keluar rumah dan tetap menerapkan protokol kesehatan yang ditetapkan pemerintah.

2.2 Analisis Regresi

Analisis regresi merupakan suatu metode yang digunakan untuk mengetahui hubungan antara suatu variabel dependen terhadap satu atau lebih variabel independen (Ghambir dkk 2020). Salah satu model regresi paling umum adalah regresi linear seperti pada rumus 2.1.

$$y = \alpha + \beta X + \varepsilon \quad (2.1)$$

dimana :

- y** = variabel dependen (variabel yang akan coba diprediksi)
- X** = variabel independen (variabel yang digunakan untuk memprediksi y)
- α** = konstanta
- β** = nilai koefisien regresi (kemiringan)
- ε** = nilai error (perbedaan nilai prediksi y dengan nilai y yang sebenarnya)

Model lain dari regresi adalah regresi polynomial seperti pada rumus 2.2. Regresi polynomial sering digunakan pada range data lebar dan tidak berbentuk linear. Regresi polynomial dapat memperkirakan hubungan antara variabel dependen dan independen dengan akurat dan mudah dikombinasikan dengan fungsi lain

$$y = \alpha_0 + \alpha_1 * X_1 + \alpha_2 * X_2^2 + \dots + \alpha_n * X_n^n \quad (2.2)$$

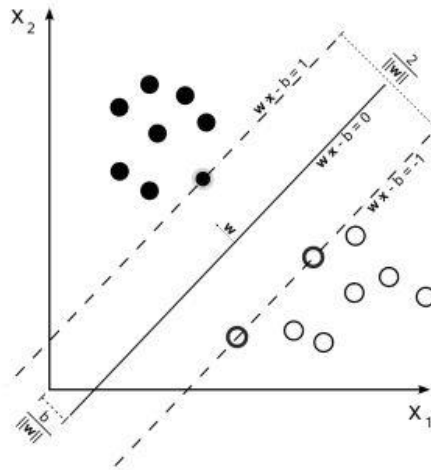
dimana :

- y** = variabel dependen (variabel yang akan coba diprediksi)
- X** = variabel independen (variabel yang digunakan untuk memprediksi y)
- α** = nilai koefisien regresi untuk data ke – n

2.3 Support Vector Machine

Support Vector Machine (SVM) adalah salah satu metode *machine learning* yang digunakan untuk memprediksi klasifikasi menggunakan konsep pencarian *hyperplane* dengan margin maksimum. Awalnya SVM terfokus untuk memecahkan masalah klasifikasi pola, namun saat ini SVM dapat diterapkan pada 7 aplikasi yang

lebih luas seperti pendekatan fungsi, estimasi regresi, dan prediksi deret waktu. Karakteristik penting dari SVM yaitu hasil proses training berupa *support vector* yang dapat digunakan kembali pada proses *testing*.



Gambar 2.1 Ilustrasi SVM (sumber: Abe,2005)

Melalui gambar di atas kita dapat melihat ada 3 garis sejajar diagonal dari kiri bawah menuju kanan atas. Saya akan membaginya menjadi 3 garis:

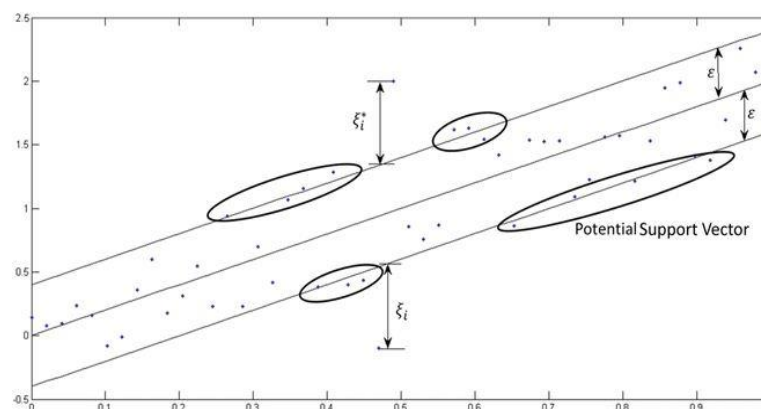
- Garis paling atas (garis putus-putus) adalah garis batas +, artinya garis ini merupakan garis yang melewati/melalui poin (satu data poin) yang menjadi bagian dari klaster +. Titik-titik (*data points*) ini sering disebut juga dengan istilah *support vectors*. Titik yang dilewati garis ini adalah titik terluar dari klaster + sekaligus menjadi batas bagi klaster +. Jarak antara garis ini dengan garis tengah haruslah tegak lurus (*perpendicular*).
- Garis tengah adalah garis pembatas antara kedua zona (+ dan -), yang dalam SVM disebut dengan istilah *hyperplane* atau *separating plane/line*. Garis ini memisahkan dua klaster, yaitu klaster + dan -. Atau bisa saja kita sebut dengan klaster A dan B, 1 dan 2, dan seterusnya (penamaan bebas, namun konvensi umum menggunakan + dan -).
- Garis paling bawah (garis putus-putus) adalah garis batas -, artinya garis ini merupakan garis yang memiliki melewati satu data poin yang menjadi bagian dari klaster -. Penjelasanannya sama dengan garis batas +.

Perlu diingat bahwa jarak antara garis paling atas dengan garis tengah, adalah sama dengan jarak antara garis paling bawah dengan garis tengah. Jarak ini kita sebut dengan epsilon dengan simbol ϵ .

Inti dari gambar di atas adalah bagaimana agar bisa membagi dua zona (dua kluster) dengan sebisa mungkin memaksimalkan jarak epsilon (ϵ) nya, sehingga didapat sebuah pembagian yang paling optimal. Hasil pembagian optimal tersebut adalah didapat dua zona yang semua data poin masuk ke dalam salah satu dari dua zona ini, dengan tingkat kepastian yang presisi (jarak epsilon paling besar). Sebagai ringkasan, SVM adalah sebuah cara membagi sebuah dataset ke dalam 2 jenis dataset dengan menggunakan sebuah *hyperplane*.

2.4 Support Vector Regression

Support Vector Regression (SVR) merupakan versi regresi dari metode *Support Vector Machine* (SVM), hasil *output* dari metode ini menghasilkan bilangan riil atau kontinu. SVR diperkenalkan oleh Drucker untuk mengatasi masalah regresi. SVR merupakan metode regresi yang mampu mengatasi model *overfitting* serta mampu menunjukkan performansi yang bagus. Tujuan dari SVR adalah untuk menemukan fungsi $f(x)$ sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang sesuai dengan semua input data dengan sebuah error ϵ dan membuat ϵ sekecil mungkin (Scholkopt dkk,2002).



Gambar 2.2 Ilustrasi SVR (sumber: Abe,2005)

Gambar di atas menunjukkan sebuah *hyperplane* (garis diagonal di tengah) yang diapit oleh dua garis batas + dan garis batas - (dijelaskan di atas). Kita juga

melihat ada ε sebagai jarak antara *hyperplane* dengan 2 garis batas tadi. Bisa dilihat ada beberapa datapoin yang dilingkari yang menjadi *potential support vectors*. Artinya titik-titik (data points) ini merupakan data poin yang bisa menjadi calon pembatas, sehingga semua data poin bisa masuk ke dalam satu klaster, dengan tetap sebisa mungkin meminimasi nilai ε nya. Sehingga jika divisualisasikan, garis *hyperplane* nya sebisa mungkin melewati semua titik-titik data (*data points*) tadi dan gambarnya akan tampak seperti grafik regresi pada umumnya.

Perbedaan SVM dan SVR adalah Jika SVM itu tujuannya membagi dataset (klasifikasi) ke dalam 2 zona, maka SVR sebaliknya, yaitu bagaimana caranya agar semua dataset masuk ke dalam satu zona, dengan tetap meminimasi nilai epsilon (ε). SVM membagi menjadi dua, sementara SVR memasukkan semuanya menjadi satu.

Misalkan ada l data *training*, $(x_i, y_i), i = 1, \dots, l$ dimana x_i merupakan *vector input* $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \subseteq R^n$ dan *output* skalar $\mathbf{y} = \{y_1, \dots, y_l\} \subseteq R$ dan l adalah banyaknya data *training*. Dengan SVR ingin ditentukan suatu fungsi $f(\mathbf{x})$ yang mempunyai deviasi paling besar ε dari target *actual* y_1 , untuk semua data *training*. Jika nilai $\varepsilon = 0$ maka diperoleh suatu persamaan regresi yang sempurna. Dengan metode SVR diperoleh fungsi :

$$f(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\varphi}(\mathbf{x}) + \mathbf{b} \quad (2.3)$$

Dengan :

\mathbf{w} = *vector* bobot berdimensi l

$\boldsymbol{\varphi}(\mathbf{x})$ = fungsi yang menentukan \mathbf{x} pada ruang dengan l dimensi

\mathbf{b} = bias (Santosa, 2007).

Banyak teknik data *mining* atau *machine learning* yang dikembangkan dengan asumsi kelinearan, sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus yang *linear*. Dengan metode kernel, suatu data \mathbf{x} di *input space* dipetakan ke *feature space* dengan dimensi yang lebih tinggi melalui $\boldsymbol{\varphi}$ (Santosa, 2007).

Macam fungsi kernel diantaranya yaitu :

1. Linier : $\mathbf{K}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
2. Polynomial : $\mathbf{K}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + \mathbf{c})^d$

3. Radial basis function (RBF) : : $K(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2)$ dengan $\gamma = \frac{1}{2\sigma^2}$
4. Sigmoid : $K(\mathbf{x}, \mathbf{y}) = \tanh(\sigma(\mathbf{x} \cdot \mathbf{y}) + c)$ dan \mathbf{x} dan \mathbf{y} adalah pasangan dua data dari semua bagian data latih. Parameter $\sigma, c, d > 0$, merupakan konstanta. Dalam fungsi kernel yang legitimate diberikan oleh Teori Mercer dimana fungsi tersebut harus memenuhi syarat kontinu dan positif definit.

2.5 Pengukuran Performa Hasil Prediksi

Performa suatu prediksi dari suatu metode diukur dari error hasil prediksi tersebut. Performa hasil prediksi pada penelitian kali ini diukur menggunakan *Mean Absolute Percentage Error* (MAPE). MAPE adalah rata-rata dari keseluruhan persentase selisih antara data aktual dan data hasil prediksi, dengan formula MAPE sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{\hat{y}_t} \right| \times 100 \quad (2.4)$$

dengan y_t adalah kasus Covid-19 aktual, \hat{y}_t adalah kasus Covid-19 prediksi, n adalah banyaknya data dan t adalah waktu.

BAB 3. METODE PENELITIAN

3.1 Data

Pada penelitian ini data yang digunakan adalah data harian kasus covid-19 di Jawa Timur yang merupakan data sekunder. Data ini dapat diakses di <https://www.kaggle.com> dengan periode 18 Maret 2020 sampai 4 Oktober 2020. Variabel yang digunakan adalah variabel prediktor dan variabel respon. Variabel respon (y) dalam penelitian ini adalah total kasus harian di Jawa Timur yang digunakan sebagai target. Variabel prediktor (x) pada penelitian ini sebagai variabel pendukung adalah kasus baru, kematian baru, sembuh baru, kasus aktif baru, total kematian, total sembuh pada tanggal 18 Maret 2020 sampai 4 Oktober 2020. Untuk data lengkapnya dapat dilihat pada tabel A dilampiran.

Tabel 3.1 Variabel Prediktor

Variabel Prediktor (X)	
X_1	Kasus Baru
X_2	Kematian Baru
X_3	Sembuh Baru
X_4	Kasus Aktif Baru
X_5	Total Kematian
X_6	Total Sembuh

3.2 Langkah – langkah Penelitian

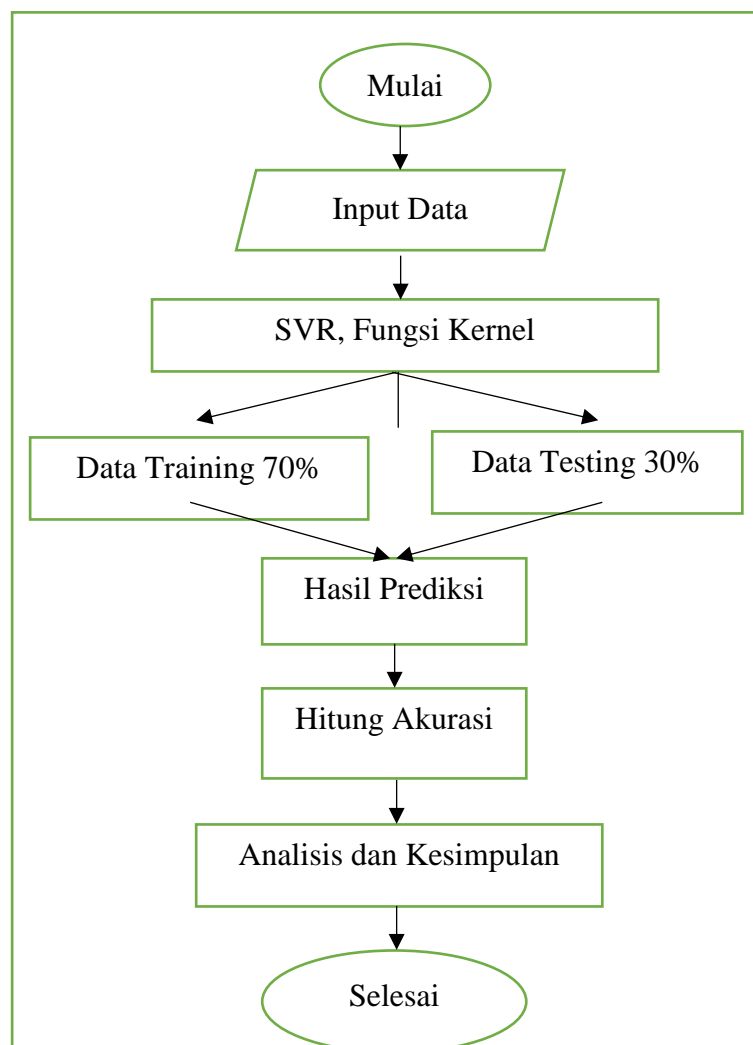
Adapun langkah –langkah yang akan dilakukan dalam penelitian ini,

1. Mengambil data dari <https://www.kaggle.com> tahun 2020–2021 kemudian diolah menggunakan *software* R Studio. Untuk mengetahui sintak nya, dapat dilihat pada bagian B dilampiran.
2. Membagi data menjadi 70% data training dari tanggal 18 Maret 2020 sampai 5 Agustus 2020 dan 30% data testing dari tanggal 6 Agustus 2020 sampai 4 Oktober 2020.

3. *Fitting* SVR dan menentukan fungsi kernel yang akan digunakan yaitu kernel linier, polynomial, radial, dan sigmoid.
4. Melakukan uji model terbaik dan menghitung *Mean Absolute Percentage Error*.
5. Melakukan prediksi kasus Covid-19 dengan model *Support Vector Regression*.
6. Membuat kesimpulan dari hasil yang diperoleh.
7. Selesai.

3.3 Diagram Alir

Berikut ini adalah alur pemetaan untuk memproses data dalam penelitian pada Gambar 3.1.



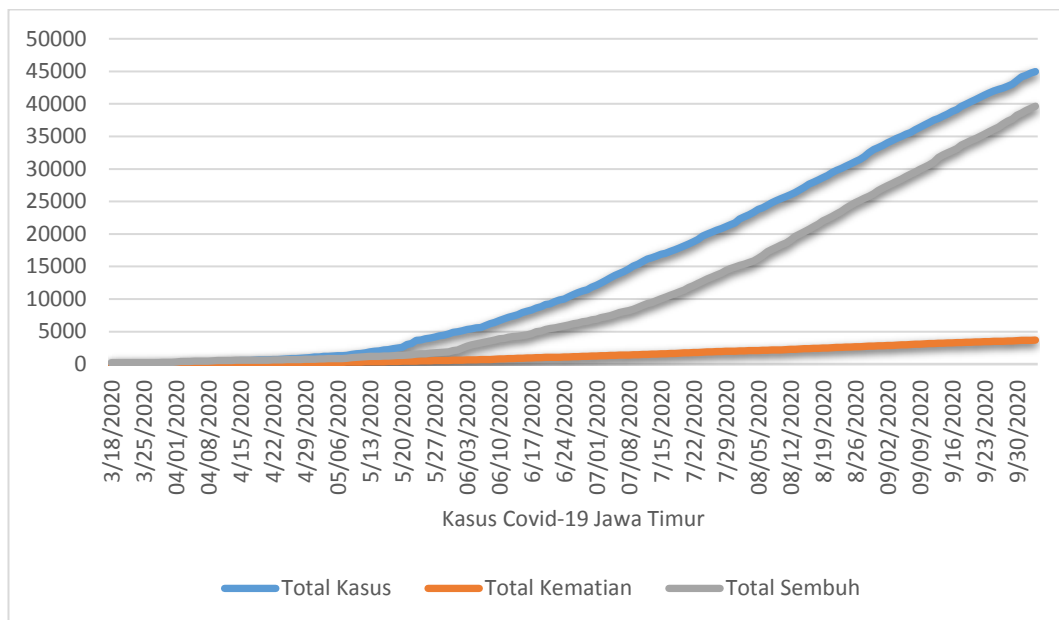
Gambar 3.1 Diagram Alir

BAB 4. HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang hasil dan pembahasan yang telah diperoleh setelah melakukan analisis sesuai dengan metode penelitian. Pertama akan dijelaskan tentang data yang digunakan dalam penelitian dan selanjutnya akan dijelaskan tentang prediksi SVR pada program R.

4.1 Deskripsi Kasus Covid-19 di Jawa Timur

Berikut ini adalah grafik kasus Covid-19 di Jawa Timur. Ditunjukkan pada Gambar 4.1



Gambar 4.1 Plot Kasus Covid-19 Jawa Timur

Secara grafis, grafik data aktual kasus covid-19 harian Provinsi Jawa Timur terus mengalami kenaikan. Pada tanggal 18 Maret 2020 sampai 27 April 2020 kenaikan tidak signifikan yakni 851 kasus. Menuju tanggal 28 April 2020 sampai 30 September 2020 terjadi kenaikan yang signifikan menjadi 43.534 kasus. Sementara kasus sembuh karena Covid-19 juga mengalami kenaikan. Pada tanggal 18 Maret 2020 sampai 29 April 2020 kenaikan tidak signifikan yakni 728 orang. Menuju tanggal 30 April 2020 sampai 30 September terjadi kenaikan yang signifikan menjadi 38.227 orang. Untuk kasus kematian mengalami kenaikan tetapi tidak signifikan sebesar 3.618 jiwa sampai tanggal 30 September 2020.

4.2 Data Penelitian

Data dalam penelitian ini menggunakan 201 objek yang meliputi 7 variabel yakni jumlah kasus baru harian, kematian baru harian, sembuh baru harian, kasus aktif baru harian, total kasus, total kematian, total sembuh dari tanggal 18 Maret 2020 sampai 22 November 2020. Variabel independen (x) pada penelitian ini adalah jumlah kasus baru harian, kematian baru harian, sembuh baru harian, kasus aktif baru harian, total kasus, total kematian, dan total sembuh. Variabel dependen (y) pada penelitian ini adalah total kasus. Pada metode SVR dilakukan pembagian data 70% *training* dan 30% *testing*. Untuk data lengkapnya dapat dilihat di lampiran A.

4.3 Prediksi dengan *Support Vector Regression (SVR)* di Jawa Timur

Prediksi SVR menggunakan program R dengan paket yang digunakan e1071. Prediksi yang dilakukan yaitu untuk mengetahui kasus Covid-19 di Jawa Timur. Data yang telah diinputkan ke dalam program R kemudian dibagi menjadi 141 data *training* dan 60 data *testing*.

4.2.1 *Tune Parameter SVR*

Tune parameter cost dilakukan untuk menguji nilai *cost* sehingga didapatkan model yang terbaik dan mencegah terjadinya kesalahan regresi. *Tune parameter cost* bernilai 0,001 , 0,01 , 0,1 , 1 , 5 , 10 , 100 pada masing-masing fungsi kernel menggunakan *training* 70%. Hasil *tuning* berupa nilai eror pada masing-masing kernel untuk tiap parameter cost ditunjukkan pada tabel 4.1.

Tabel 4.1 *Tune eror* dengan 7 parameter *cost*

Parameter cost	Kernel radial	Kernel linear	Kernel sigmoid	Kernel polynomial
0,001	59.135.817,6	17.954.402,3	55.461.760	47.476.369
0,01	39.454.364,3	726.968,2	19.550.470	12.142.781
0,1	4.847.246,6	140.286,2	10.542.860	12.377.002
1	984.767,2	137.732,2	381.094.100	4.630.304
5	950.052,3	137.732,2	8.453.459.000	4.933.441
10	950.052,3	137.732,2	33.360.730.000	5.459.188
100	950.052,3	137.732,2	3.203.718.000.000	13.547.243

Dari Tabel 4.1 didapatkan *cost parameter* terbaik kernel linier dengan nilai 137.732,2 yaitu 0,1. Hasil *tuning* dengan *parameter cost* terbaik digunakan untuk menguji akurasi data testing. Jika kita tambahkan nilai parameter cost nya ke 1000, 10000 dan seterusnya, maka hasil error nya konvergen dengan cost 0,1. Untuk data lengkapnya dapat dilihat pada lampiran C.

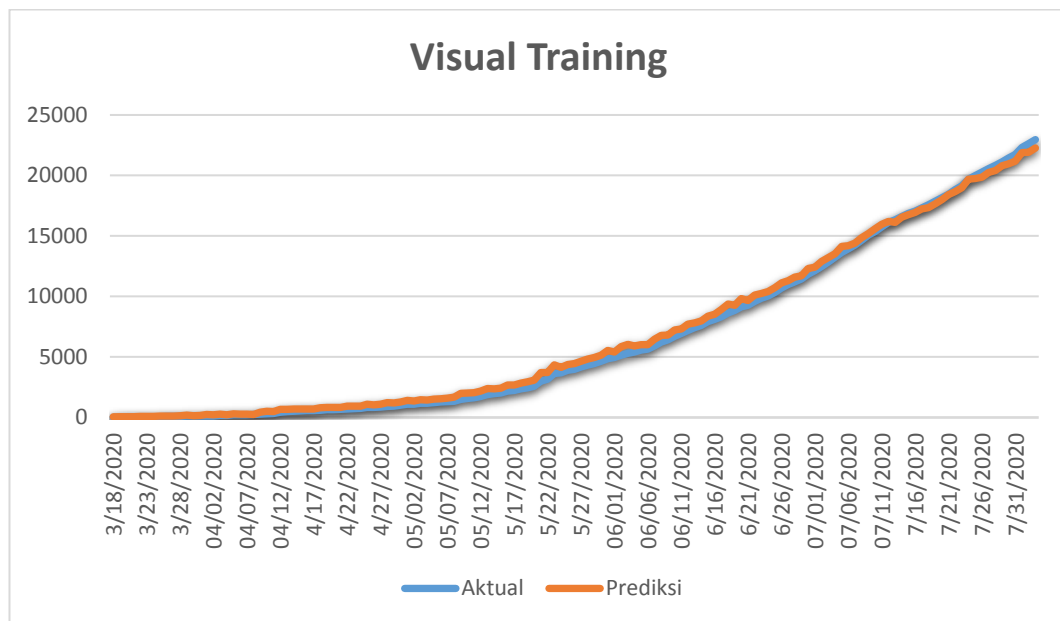
4.2.2 Penentuan Fungsi Kernel Sebagai Parameter Terbaik

Penentuan fungsi kernel dilakukan percobaan data *training* dan akan dipilih 4 tipe kernel yaitu *radial*, *linier*, *sigmoid*, dan *polynomial*. Fungsi kernel terbaik merupakan tipe kernel yang mempunyai nilai MAPE terkecil. Hasil optimasi untuk penentuan tipe kernel dengan menggunakan data *training* sebesar 70% dapat dilihat pada Tabel 4.2.

Tabel 4.2 Penentuan tipe kernel terbaik data *training*

Tipe Kernel	MAPE
Radial	150,5
Linier	22,4
Sigmoid	2507,8
Polynomial	140,9

Pada Tabel 4.2, dapat dilihat bahwa tipe kernel dengan MAPE terkecil adalah kernel linier yaitu 22,4%. Dengan menggunakan tipe kernel linier dan nilai cost parameter 0,1 akan ditentukan parameter terbaik dengan menggunakan data *testing*. Parameter terbaik yakni parameter dengan nilai MAPE terkecil. Penentuan parameter terbaik pada prediksi kali ini dilakukan dengan memasukkan beberapa input nilai parameter cost dan MAPE training.



Gambar 4.2 Plot Data Training

Secara grafis, grafik data aktual kasus covid-19 harian Provinsi Jawa Timur dan hasil prediksi dari model SVR periode *training* dapat dilihat pada gambar 4.2. Hasil prediksi yang didapat berdasarkan tabel menunjukkan bahwa metode SVR dapat memprediksi kasus Covid-19 harian di Provinsi Jawa Timur dengan baik karena hasil prediksi hampir mendekati data aktual. Pada gambar 4.2 juga memiliki nilai fluktuatif dari nilai riil dan nilai prediksi data kasus covid-19 Provinsi Jawa Timur. Pada tanggal 18 Maret 2020 sampai 2 Mei 2020 baik data aktual maupun data prediksi mengalami kenaikan. Kenaikan secara signifikan terjadi pada tanggal 3 Mei 2020 sampai puncaknya pada tanggal 31 Juli 2020. Tahap selanjutnya adalah memvalidasi model SVR dengan menggunakan data *testing*. Pada penelitian ini, validasi model menggunakan data kasus Covid-19 harian Provinsi Jawa Timur pada tanggal 6 Agustus 2020 sampai 4 Oktober 2020. Hal ini bertujuan untuk melihat akurasi prediksi model SVR pada periode testing.

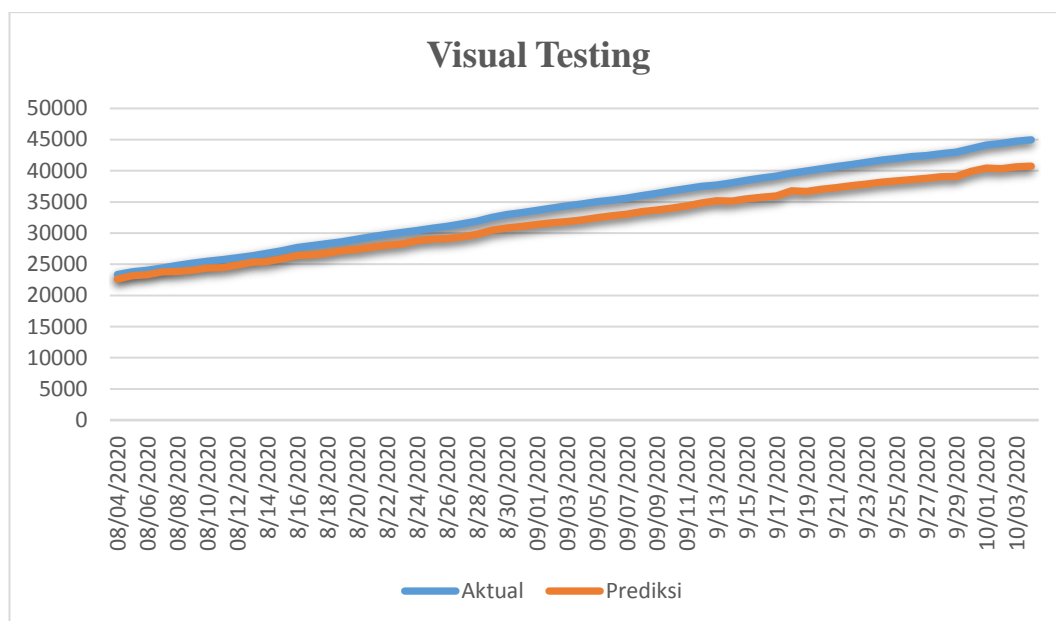
4.2.3 Prediksi Kasus Covid-19 dengan Model SVR

Parameter terbaik yang telah diperoleh pada data *training* selanjutnya diterapkan pada data *testing*. Input yang digunakan dalam prediksi tertera pada Tabel 4.3.

Tabel 4.3 Input prediksi SVR

Deskripsi	Nilai
Tipe Kernel	Linier
Nilai <i>Cost</i>	0,1

Hasil prediksi dengan metode SVR yang dilakukan dengan perbandingan 70% *training* 30% *testing* dengan *input* tabel diatas menghasilkan nilai *MAPE* 6,7%. Hasil tersebut termasuk pada kategori performa baik. Untuk data lengkapnya dapat dilihat pada lampiran E. Visualisasi hasil prediksi dapat dilihat pada Gambar 4.3.

Gambar 4.3 Plot Data *Testing*

Secara visual perbandingan hasil prediksi dapat terlihat pada Gambar 4.3. Model ini terlihat memiliki pola mengikuti data aktual yakni mengalami kenaikan. Pada tanggal 4 Agustus 2020 sampai 17 Agustus 2020 pola grafik prediksi mampu mengikuti pola aktual dengan baik. Sementara untuk tanggal 18 Agustus 2020 sampai 3 Oktober 2020 baik data aktual maupun data prediksi mengalami kenaikan tetapi tidak mampu memprediksi dengan baik jika dilihat dari jumlah kasus perharinya pada tanggal tersebut. Maka dapat diperoleh hasil prediksi total kasus Covid-19 harian di Provinsi Jawa Timur. Untuk data lengkapnya dapat dilihat pada lampiran D.

BAB 5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis data dan pembahasan maka diperoleh kesimpulan sebagai berikut :

1. Penggunaan metode SVR untuk memprediksi kasus Covid-19 di Jawa Timur didapatkan hasil bahwa prediksi terbaik adalah ketika menggunakan kernel linier dengan dukungan nilai parameter cost dan nilai MAPE.
2. Hasil prediksi kasus Covid-19 di Jawa Timur dengan metode SVR menggunakan rasio perbandingan data *training* dan *testing* sebesar 70:30 menghasilkan nilai *MAPE* 6,7% Hasil tersebut termasuk pada kategori performa yang baik.

5.2 Saran

1. Melakukan tuning parameter yang lebih bervariasi, misalnya dengan mencoba library lain untuk melakukan tuning parameter sehingga hasil dapat dibandingkan dan dapat memberikan performa yang lebih baik.
2. Menambah jumlah nilai parameter cost untuk setiap kernel sehingga memungkinkan mendapat nilai lebih minimal nilai cost pada setiap kernel.

DAFTAR PUSTAKA

- Abe, S. 2005. *Support Vector Machine for Pattern Classification*. Springer - Verlag. London Limited.
- Ghambir., dkk 2020. Regression Analysis of COVID-19 using Machine Learning Algorithms. IEEE. (CFP20V90-ART): 65–71.
- Gennaro, F. 2020. Coronavirus Diseases (COVID-19) Current Status and Future Perspectives. *A Narrative Review. International Journal of Environmental Research and Public Health*. 17(2690): 1–11.
- Kemkes RI. 2020. *Situasi Terkini Perkembangan Novel Coronavirus (Covid-19)*. <https://covid19.kemkes.go.id/situasi-infeksi-emerging/info-corona-virus/situasi-terkini-perkembangan-coronavirus-disease-covid-19-31-agustus2020/#.X0y2bNwzZxQ> [Diakses 31 Agustus 2020].
- Ray, S. 2015. *7 Regression Techniques you should know*. <https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guideregression/> [Diakses 29 Januari 2021].
- Santosa,Budi. 2000. *Tutorial Support Vector Machin*. ITS,Surabaya
- Scholkopt B., dkk. 2002. *Learning With Kernel*. MIT Press.
- Susilo,A. 2020. Coronavirus Disease 2019. Tinjauan Literatur Terkini. *Jurnal Penyakit Dalam Indonesia*. 7(1): 45–67.
- Tukur, K. dan Usman, A. U. 2018. Binary Logistic Regression Analysis on Admitting Students Using Jamb Score. *International Journal of Current Research*. 8(1): 25235–25239.
- World Health Organization. 2020. *World Health Organization Coronavirus Disease 2019 Global Situation 2020*. <https://covid19.who.int/>. [Diakses 31 Agustus 2020].
- Xu, X. 2020. Evolution of Novel Coronavirus from The Ongoing Wuhan Outbreak and Modeling of Its Spike Protein For Risk Of Human Transmission. *Science China Life Sciences*. 63(3): 457–460.
- Zhu, N. 2020. A Novel Coronavirus from Patients with Pneumonia in China 2019. *The New England Journal of Medicine*. 382(8): 727–733.

LAMPIRAN

Lampiran A. Data Aktual Kasus Covid-19 di Jawa Timur

Tanggal	Kasus Baru	Kematian Baru	Sembuh Baru	Kasus Aktif Baru	Total Kematian	Total Sembuh	Total Kasus
3/18/2020	7	1	0	6	33	215	18
3/19/2020	3	0	0	3	33	215	21
3/20/2020	9	0	0	9	33	215	30
3/21/2020	9	1	0	8	34	215	39
3/22/2020	14	0	0	14	34	215	53
3/23/2020	2	1	0	1	35	215	55
3/24/2020	3	1	0	2	36	215	58
3/25/2020	9	0	1	8	36	216	67
3/26/2020	13	1	0	12	37	216	80
3/27/2020	3	1	0	2	38	216	83
3/28/2020	14	2	2	10	40	218	97
3/29/2020	12	1	28	-17	41	246	109
3/30/2020	7	2	0	5	43	246	116
3/31/2020	1	0	0	1	43	246	117
04/01/2020	9	4	77	-72	47	323	126
04/02/2020	0	12	91	-103	59	414	126
04/03/2020	47	1	7	39	60	421	173
04/04/2020	5	2	3	0	62	424	178
04/05/2020	35	2	8	25	64	432	213
04/06/2020	9	5	11	-7	69	443	222
04/07/2020	5	5	2	-2	74	445	227
04/08/2020	15	23	15	-23	97	460	242
04/09/2020	36	12	43	-19	109	503	278
04/10/2020	30	4	34	-8	113	537	308
04/11/2020	15	9	27	-21	122	564	323
04/12/2020	111	1	7	103	123	571	434
4/13/2020	52	3	6	43	126	577	486
4/14/2020	36	2	10	24	128	587	522
4/15/2020	25	4	4	17	132	591	547
4/16/2020	13	4	3	6	136	594	560
4/17/2020	10	6	3	1	142	597	570
4/18/2020	34	3	7	24	145	604	604
4/19/2020	43	6	2	35	151	606	647
4/20/2020	5	3	1	1	154	607	652
4/21/2020	10	5	2	3	159	609	662
4/22/2020	36	5	30	1	164	639	698
4/23/2020	27	4	12	11	168	651	725

4/24/2020	20	11	14	-5	179	665	745
4/25/2020	75	7	19	49	186	684	820
4/26/2020	19	8	4	7	194	688	839
4/27/2020	12	5	9	-2	199	697	851
4/28/2020	58	1	21	36	200	718	909
4/29/2020	20	3	10	7	203	728	929
4/30/2020	75	5	8	62	208	736	1004
05/01/2020	83	5	33	45	213	769	1087
05/02/2020	7	10	46	-49	223	815	1094
05/03/2020	77	6	9	62	229	824	1171
05/04/2020	13	8	7	-2	237	831	1184
05/05/2020	41	7	21	13	244	852	1225
05/06/2020	48	13	11	24	257	863	1273
05/07/2020	34	14	9	11	271	872	1307
05/08/2020	30	19	46	-35	290	918	1337
05/09/2020	158	12	63	83	302	981	1495
05/10/2020	78	7	48	23	309	1029	1573
05/11/2020	54	13	70	-29	322	1099	1627
05/12/2020	111	11	28	72	333	1127	1738
5/13/2020	141	7	47	87	340	1174	1879
5/14/2020	85	14	17	54	354	1191	1964
5/15/2020	57	13	12	32	367	1203	2021
5/16/2020	168	11	24	133	378	1227	2189
5/17/2020	57	11	30	16	389	1257	2246
5/18/2020	124	13	15	96	402	1272	2370
5/19/2020	84	13	48	23	415	1320	2454
5/20/2020	124	19	43	62	434	1363	2578
5/21/2020	431	14	12	405	448	1375	3009
5/22/2020	155	9	51	95	457	1426	3164
5/23/2020	465	20	90	355	477	1516	3629
5/24/2020	72	11	55	6	488	1571	3701
5/25/2020	200	16	34	150	504	1605	3901
5/26/2020	86	14	67	5	518	1672	3987
5/27/2020	166	15	44	107	533	1716	4153
5/28/2020	168	13	55	100	546	1771	4321
5/29/2020	132	8	52	72	554	1823	4453
5/30/2020	180	12	54	114	566	1877	4633
5/31/2020	232	16	203	13	582	2080	4865
06/01/2020	95	16	96	-17	598	2176	4960
06/02/2020	171	19	298	-146	617	2474	5131
06/03/2020	173	13	304	-144	630	2778	5304
06/04/2020	107	15	179	-87	645	2957	5411

06/05/2020	144	22	172	-50	667	3129	5555
06/06/2020	84	19	117	-52	686	3246	5639
06/07/2020	270	18	147	105	704	3393	5909
06/08/2020	319	20	161	138	724	3554	6228
06/09/2020	189	24	127	38	748	3681	6417
06/10/2020	275	29	203	43	777	3884	6692
06/11/2020	251	32	88	131	809	3972	6943
06/12/2020	245	29	185	31	838	4157	7188
6/13/2020	218	23	93	102	861	4250	7406
6/14/2020	194	22	63	109	883	4313	7600
6/15/2020	312	30	95	187	913	4408	7912
6/16/2020	204	22	59	123	935	4467	8116
6/17/2020	210	27	239	-56	962	4706	8326
6/18/2020	319	23	269	27	985	4975	8645
6/19/2020	196	22	115	59	1007	5090	8841
6/20/2020	300	18	257	25	1025	5347	9141
6/21/2020	131	16	126	-11	1041	5473	9272
6/22/2020	329	11	116	202	1052	5589	9601
6/23/2020	241	17	130	94	1069	5719	9842
6/24/2020	183	19	161	3	1088	5880	10025
6/25/2020	304	23	135	146	1111	6015	10329
6/26/2020	361	32	186	143	1143	6201	10690
6/27/2020	287	25	105	157	1168	6306	10977
6/28/2020	236	21	182	33	1189	6488	11213
6/29/2020	226	21	102	103	1210	6590	11439
6/30/2020	394	24	198	172	1234	6788	11833
07/01/2020	305	25	116	164	1259	6904	12138
07/02/2020	342	28	222	92	1287	7126	12480
07/03/2020	372	20	160	192	1307	7286	12852
07/04/2020	384	31	154	199	1338	7440	13236
07/05/2020	411	14	262	135	1352	7702	13647
07/06/2020	280	18	219	43	1370	7921	13927
07/07/2020	321	21	153	147	1391	8074	14248
07/08/2020	393	17	164	212	1408	8238	14641
07/09/2020	416	27	180	209	1435	8418	15057
07/10/2020	295	17	311	-33	1452	8729	15352
07/11/2020	393	19	310	64	1471	9039	15745
07/12/2020	388	21	261	106	1492	9300	16133
7/13/2020	197	22	165	10	1514	9465	16330
7/14/2020	278	26	272	-20	1540	9737	16608
7/15/2020	261	39	315	-93	1579	10052	16869
7/16/2020	191	17	254	-80	1596	10306	17060

7/17/2020	249	21	306	-78	1617	10612	17309
7/18/2020	257	26	224	7	1643	10836	17566
7/19/2020	302	44	285	-27	1687	11121	17868
7/20/2020	287	26	257	4	1713	11378	18155
7/21/2020	305	25	370	-90	1738	11748	18460
7/22/2020	359	32	311	16	1770	12059	18819
7/23/2020	323	37	363	-77	1807	12422	19142
7/24/2020	563	30	348	185	1837	12770	19705
7/25/2020	270	22	353	-105	1859	13123	19975
7/26/2020	302	33	260	9	1892	13383	20277
7/27/2020	286	18	309	-41	1910	13692	20563
7/28/2020	273	31	311	-69	1941	14003	20836
7/29/2020	286	19	366	-99	1960	14369	21122
7/30/2020	315	20	283	12	1980	14652	21437
7/31/2020	294	22	261	11	2002	14913	21731
08/01/2020	581	19	240	322	2021	15153	22312
08/02/2020	330	23	220	87	2044	15373	22642
08/03/2020	321	17	258	46	2061	15631	22963
08/04/2020	391	19	247	125	2080	15878	23354
08/05/2020	414	11	347	56	2091	16225	23768
08/06/2020	265	19	490	-244	2110	16715	24033
08/07/2020	365	21	576	-232	2131	17291	24398
08/08/2020	402	24	345	33	2155	17636	24800
08/09/2020	367	29	317	21	2184	17953	25167
08/10/2020	298	18	382	-102	2202	18335	25465
08/11/2020	288	23	277	-12	2225	18612	25753
08/12/2020	298	34	450	-186	2259	19062	26051
8/13/2020	327	36	569	-278	2295	19631	26378
8/14/2020	402	41	359	2	2336	19990	26780
8/15/2020	416	27	366	23	2363	20356	27196
8/16/2020	471	23	389	59	2386	20745	27667
8/17/2020	313	27	389	-103	2413	21134	27980
8/18/2020	318	24	363	-69	2437	21497	28298
8/19/2020	338	27	485	-174	2464	21982	28636
8/20/2020	366	25	367	-26	2489	22349	29002
8/21/2020	451	39	367	45	2528	22716	29453
8/22/2020	343	30	382	-69	2558	23098	29796
8/23/2020	282	28	369	-115	2586	23467	30078
8/24/2020	340	22	491	-173	2608	23958	30418
8/25/2020	364	20	460	-116	2628	24418	30782
8/26/2020	322	31	398	-107	2659	24816	31104
8/27/2020	362	31	353	-22	2690	25169	31466

8/28/2020	424	32	337	55	2722	25506	31890
8/29/2020	615	24	264	327	2746	25770	32505
8/30/2020	458	34	398	26	2780	26168	32963
8/31/2020	315	38	493	-216	2818	26661	33278
09/01/2020	351	19	405	-73	2837	27066	33629
09/02/2020	359	17	360	-18	2854	27426	33988
09/03/2020	371	34	331	6	2888	27757	34359
09/04/2020	330	43	347	-60	2931	28104	34689
09/05/2020	322	26	336	-40	2957	28440	35011
09/06/2020	305	32	391	-118	2989	28831	35316
09/07/2020	290	26	376	-112	3015	29207	35606
09/08/2020	425	27	348	50	3042	29555	36031
09/09/2020	356	38	391	-73	3080	29946	36387
09/10/2020	370	27	315	28	3107	30261	36757
09/11/2020	369	34	359	-24	3141	30620	37126
09/12/2020	364	24	442	-102	3165	31062	37490
9/13/2020	236	30	710	-504	3195	31772	37726
9/14/2020	358	31	346	-19	3226	32118	38084
9/15/2020	371	24	346	1	3250	32464	38455
9/16/2020	358	24	340	-6	3274	32804	38813
9/17/2020	323	29	333	-39	3303	33137	39136
9/18/2020	488	23	538	-73	3326	33675	39624
9/19/2020	351	31	353	-33	3357	34028	39975
9/20/2020	350	21	321	8	3378	34349	40325
9/21/2020	336	16	286	34	3394	34635	40661
9/22/2020	336	22	379	-65	3416	35014	40997
9/23/2020	342	23	335	-16	3439	35349	41339
9/24/2020	359	24	363	-28	3463	35712	41698
9/25/2020	284	22	374	-112	3485	36086	41982
9/26/2020	269	17	344	-92	3502	36430	42251
9/27/2020	197	19	460	-282	3521	36890	42448
9/28/2020	280	22	451	-193	3543	37341	42728
9/29/2020	275	24	321	-70	3567	37662	43003
9/30/2020	531	51	565	-85	3618	38227	43534
10/01/2020	601	22	372	207	3640	38599	44135
10/02/2020	269	22	363	-116	3662	38962	44404
10/03/2020	319	15	343	-39	3677	39305	44723
10/04/2020	238	18	331	-111	3695	39636	44961

**www.kaggle.com*

Lampiran B1. Prediksi dengan Kernel Radial

```
#Memanggil data
data= read.csv("G:/ProjectCovid/kasus covid 19 jatim.csv",sep = ";")
data=data[-1]
i <- c(1:8)
data[ , i] <- apply(data[ , i], 2,function(x)
as.numeric(as.character(x)))
sapply(data, class)
head(data)
training_set = data[1:141,1:8]
test_set = data[142:201,1:8]
#Fitting kernel SVM
library(e1071)
classifier = svm(formula = Total.Kasus ~ .,
                 data = training_set,
                 type = 'eps-regression',
                 kernel = 'radial')
summary(classifier)
#Prediction training data result
y_train_pred = predict(classifier, newdata=training_set[-8])
Mape_train=mean(abs((training_set$Total.Kasus-
y_train_pred)/training_set$Total.Kasus)) * 100
Mape_train
#mencari cost error terbaik
tuning <- tune(svm,Total.Kasus~.,data=training_set,type = 'eps-
regression',
              kernel='radial',
              ranges=list(cost=c(0.001, 0.01, 0.1, 1, 5, 10, 100)))
summary(tuning)
#Testing SVM
y_test_pred = predict(classifier, newdata=test_set[-8])
Mape_test=mean(abs((test_set$Total.Kasus-
y_test_pred)/test_set$Total.Kasus)) * 100
Mape_test
write.csv2(y_train_pred, "G:/ProjectCovid//predtrain.csv")
write.csv2(y_test_pred, "G:/ProjectCovid//predtest.csv")
```

Lampiran B2. Prediksi dengan Kernel Linier

```
#Memanggil data
data= read.csv("G:/ProjectCovid/kasus covid 19 jatim.csv",sep = ";")
data=data[-1]
i <- c(1:8)
data[ , i] <- apply(data[ , i], 2,function(x)
as.numeric(as.character(x)))
sapply(data, class)
head(data)
training_set = data[1:141,1:8]
test_set = data[142:201,1:8]
#Fitting kernel SVM
library(e1071)
classifier = svm(formula = Total.Kasus ~ .,
                 data = training_set,
                 type = 'eps-regression',
                 kernel = 'linear')
summary(classifier)
#Prediction training data result
y_train_pred = predict(classifier, newdata=training_set[-8])
Mape_train=mean(abs((training_set$Total.Kasus-
y_train_pred)/training_set$Total.Kasus)) * 100
Mape_train
#mencari cost error terbaik
tuning <- tune(svm,Total.Kasus~.,data=training_set,type = 'eps-
regression',
              kernel='linear',
              ranges=list(cost=c(0.001, 0.01, 0.1, 1, 5, 10, 100)))
summary(tuning)
#Testing SVM
y_test_pred = predict(classifier, newdata=test_set[-8])
Mape_test=mean(abs((test_set$Total.Kasus-
y_test_pred)/test_set$Total.Kasus)) * 100
Mape_test
write.csv2(y_train_pred, "G:/ProjectCovid//predtrain.csv")
write.csv2(y_test_pred, "G:/ProjectCovid//predtest.csv")
```

Lampiran B3. Prediksi dengan Kernel Polynomial

```
#Memanggil data
data= read.csv("G:/ProjectCovid/kasus covid 19 jatim.csv",sep = ";")
data=data[-1]
i <- c(1:8)
data[ , i] <- apply(data[ , i], 2,function(x)
as.numeric(as.character(x)))
sapply(data, class)
head(data)
training_set = data[1:141,1:8]
test_set = data[142:201,1:8]
#Fitting kernel SVM
library(e1071)
classifier = svm(formula = Total.Kasus ~ .,
                 data = training_set,
                 type = 'eps-regression',
                 kernel = 'polynomial')
summary(classifier)
#Prediction training data result
y_train_pred = predict(classifier, newdata=training_set[-8])
Mape_train=mean(abs((training_set$Total.Kasus-
y_train_pred)/training_set$Total.Kasus)) * 100
Mape_train
#mencari cost error terbaik
tuning <- tune(svm,Total.Kasus~.,data=training_set,type = 'eps-
regression',
              kernel='polynomial',
              ranges=list(cost=c(0.001, 0.01, 0.1, 1, 5, 10, 100)))
summary(tuning)
#Testing SVM
y_test_pred = predict(classifier, newdata=test_set[-8])
Mape_test=mean(abs((test_set$Total.Kasus-
y_test_pred)/test_set$Total.Kasus)) * 100
Mape_test
write.csv2(y_train_pred, "G:/ProjectCovid//predtrain.csv")
write.csv2(y_test_pred, "G:/ProjectCovid//predtest.csv")
```


Lampiran B4. Prediksi dengan Kernel Sigmoid

```
#Memanggil data
data= read.csv("G:/ProjectCovid/kasus covid 19 jatim.csv",sep = ";")
data=data[-1]
i <- c(1:8)
data[ , i] <- apply(data[ , i], 2,function(x)
as.numeric(as.character(x)))
sapply(data, class)
head(data)
training_set = data[1:141,1:8]
test_set = data[142:201,1:8]
#Fitting kernel SVM
library(e1071)
classifier = svm(formula = Total.Kasus ~ .,
                 data = training_set,
                 type = 'eps-regression',
                 kernel = 'sigmoid')
summary(classifier)
#Prediction training data result
y_train_pred = predict(classifier, newdata=training_set[-8])
Mape_train=mean(abs((training_set$Total.Kasus-
y_train_pred)/training_set$Total.Kasus)) * 100
Mape_train
#mencari cost error terbaik
tuning <- tune(svm,Total.Kasus~.,data=training_set,type = 'eps-
regression',
              kernel='sigmoid',
              ranges=list(cost=c(0.001, 0.01, 0.1, 1, 5, 10, 100)))
summary(tuning)
#Testing SVM
y_test_pred = predict(classifier, newdata=test_set[-8])
Mape_test=mean(abs((test_set$Total.Kasus-
y_test_pred)/test_set$Total.Kasus)) * 100
Mape_test
write.csv2(y_train_pred, "G:/ProjectCovid//predtrain.csv")
write.csv2(y_test_pred, "G:/ProjectCovid//predtest.csv")
```

Lampiran C1. Tune Parameter Radial

Best parameters cost : 5

Best performance : 950052.3

Detailed performance results:

Cost	error
1 1e-03	59135817.6
2 1e-02	39454364.3
3 1e-01	4847246.6
4 1e+00	984767.2
5 5e+00	950052.3
6 1e+01	950052.3
7 1e+02	950052.3

Lampiran C2. Tune Parameter Linier

Best parameters cost : 1

Best performance : 137732.2

Detailed performance results:

Cost	error
1 1e-03	17954402.3
2 1e-02	726968.2
3 1e-01	140286.2
4 1e+00	137732.2
5 5e+00	137732.2
6 1e+01	137732.2
7 1e+02	137732.2

Lampiran C3. Tune Parameter Polynomial

Best parameters cost : 1

Best performance: 4630304

Detailed performance results:

Cost	error
1 1e-03	47476369
2 1e-02	12142781
3 1e-01	12377002
4 1e+00	4630304
5 5e+00	4933441
6 1e+01	5459188
7 1e+02	13547243

Lampiran C4. Tune Parameter Sigmoid

Best parameters cost : 0.1

Best performance : 10542864

Detailed performance results:

Cost	error
1 1e-03	5.546176e+07
2 1e-02	1.955047e+07
3 1e-01	1.054286e+07
4 1e+00	3.810941e+08
5 5e+00	8.453459e+09
6 1e+01	3.336073e+10
7 1e+02	3.203718e+12

Lampiran D. Hasil Prediksi

Tanggal	Total Kasus (Aktual)	Prediksi
3/18/2020	18	55,70182
3/19/2020	21	62,09084
3/20/2020	30	74,28303
3/21/2020	39	77,19578

3/22/2020	53	99,83319
3/23/2020	55	87,33715
3/24/2020	58	93,33568
3/25/2020	67	113,8713
3/26/2020	80	121,9445
3/27/2020	83	119,0541
3/28/2020	97	140,018
3/29/2020	109	188,1962
3/30/2020	116	150,7776
3/31/2020	117	161,092
4/1/2020	126	247,262
4/2/2020	126	219,1223
4/3/2020	173	263,7381
4/4/2020	178	226,3493
4/5/2020	213	289,988
4/6/2020	222	270,6969
4/7/2020	227	272,9036
4/8/2020	242	242,0575
4/9/2020	278	430,4918
4/10/2020	308	500,8201
4/11/2020	323	475,9261
4/12/2020	434	676,3286
4/13/2020	486	662,6316
4/14/2020	522	695,4202
4/15/2020	547	694,8757
4/16/2020	560	704,9231
4/17/2020	570	712,2694
4/18/2020	604	793,3366
4/19/2020	647	823,6202
4/20/2020	652	825,2392
4/21/2020	662	837,8092
4/22/2020	698	934,3336
4/23/2020	725	941,2017
4/24/2020	745	932,168
4/25/2020	820	1089,23
4/26/2020	839	1053,548
4/27/2020	851	1098,102
4/28/2020	909	1226,583
4/29/2020	929	1189,724
4/30/2020	1004	1290,734
5/1/2020	1087	1406,558
5/2/2020	1094	1352,15

5/3/2020	1171	1465,779
5/4/2020	1184	1428,43
5/5/2020	1225	1526,101
5/6/2020	1273	1548,354
5/7/2020	1307	1591,842
5/8/2020	1337	1671,125
5/9/2020	1495	1998,868
5/10/2020	1573	2024,935
5/11/2020	1627	2059,994
5/12/2020	1738	2180,397
5/13/2020	1879	2385,168
5/14/2020	1964	2354,561
5/15/2020	2021	2412,602
5/16/2020	2189	2695,946
5/17/2020	2246	2685,966
5/18/2020	2370	2840,07
5/19/2020	2454	2948,061
5/20/2020	2578	3076,822
5/21/2020	3009	3707,45
5/22/2020	3164	3715,08
5/23/2020	3629	4358,124
5/24/2020	3701	4138,152
5/25/2020	3901	4379,104
5/26/2020	3987	4442,867
5/27/2020	4153	4638,91
5/28/2020	4321	4833,181
5/29/2020	4453	4956,974
5/30/2020	4633	5140,569
5/31/2020	4865	5552,447
6/1/2020	4960	5404,289
6/2/2020	5131	5856,23
6/3/2020	5304	6035,304
6/4/2020	5411	5899,415
6/5/2020	5555	6019,062
6/6/2020	5639	6019,878
6/7/2020	5909	6462,728
6/8/2020	6228	6795,123
6/9/2020	6417	6814,939
6/10/2020	6692	7220,347
6/11/2020	6943	7303,794
6/12/2020	7188	7698,802
6/13/2020	7406	7822,671

6/14/2020	7600	7978,563
6/15/2020	7912	8380,378
6/16/2020	8116	8519,413
6/17/2020	8326	8934,601
6/18/2020	8645	9370,213
6/19/2020	8841	9273,478
6/20/2020	9141	9827,429
6/21/2020	9272	9660,101
6/22/2020	9601	10122
6/23/2020	9842	10248,26
6/24/2020	10025	10403,18
6/25/2020	10329	10726,28
6/26/2020	10690	11127,38
6/27/2020	10977	11294,94
6/28/2020	11213	11605,25
6/29/2020	11439	11715,4
6/30/2020	11833	12312,35
7/1/2020	12138	12421,92
7/2/2020	12480	12890,92
7/3/2020	12852	13221,59
7/4/2020	13236	13517,18
7/5/2020	13647	14135,52
7/6/2020	13927	14185,05
7/7/2020	14248	14402,64
7/8/2020	14641	14844,31
7/9/2020	15057	15196,18
7/10/2020	15352	15580,88
7/11/2020	15745	15965,76
7/12/2020	16133	16212,83
7/13/2020	16330	16116,8
7/14/2020	16608	16549,53
7/15/2020	16869	16767,66
7/16/2020	17060	16947,78
7/17/2020	17309	17245,44
7/18/2020	17566	17350,37
7/19/2020	17868	17658,92
7/20/2020	18155	18001,39
7/21/2020	18460	18430,13
7/22/2020	18819	18673,86
7/23/2020	19142	18983,2
7/24/2020	19705	19691,27
7/25/2020	19975	19740,68

7/26/2020	20277	19857,21
7/27/2020	20563	20248,83
7/28/2020	20836	20403,94
7/29/2020	21122	20798,78
7/30/2020	21437	20969,12
7/31/2020	21731	21165,36
8/1/2020	22312	21877,14
8/2/2020	22642	21914,76
8/3/2020	22963	22267,06
8/4/2020	23354	22624,77
8/5/2020	23768	23144,76
8/6/2020	24033	23345,06
8/7/2020	24398	23786,3
8/8/2020	24800	23822,75
8/9/2020	25167	24047,62
8/10/2020	25465	24386,77
8/11/2020	25753	24454,71
8/12/2020	26051	24876,75
8/13/2020	26378	25315,99
8/14/2020	26780	25441,14
8/15/2020	27196	25915,03
8/16/2020	27667	26401,78
8/17/2020	27980	26510,35
8/18/2020	28298	26766,28
8/19/2020	28636	27193,31
8/20/2020	29002	27379,01
8/21/2020	29453	27761,36
8/22/2020	29796	28057,49
8/23/2020	30078	28250,89
8/24/2020	30418	28760,1
8/25/2020	30782	29028,01
8/26/2020	31104	29114,11
8/27/2020	31466	29408,18
8/28/2020	31890	29807,58
8/29/2020	32505	30445,66
8/30/2020	32963	30827,01
8/31/2020	33278	31089,15
9/1/2020	33629	31408,34
9/2/2020	33988	31648,9
9/3/2020	34359	31838,99
9/4/2020	34689	32092,32
9/5/2020	35011	32471,44

9/6/2020	35316	32759,92
9/7/2020	35606	33016,39
9/8/2020	36031	33444,03
9/9/2020	36387	33696,66
9/10/2020	36757	34009,09
9/11/2020	37126	34351,64
9/12/2020	37490	34819,74
9/13/2020	37726	35188,36
9/14/2020	38084	35113,31
9/15/2020	38455	35485,19
9/16/2020	38813	35768,23
9/17/2020	39136	35982,6
9/18/2020	39624	36804,86
9/19/2020	39975	36702,53
9/20/2020	40325	37020,66
9/21/2020	40661	37269,08
9/22/2020	40997	37621,92
9/23/2020	41339	37847,4
9/24/2020	41698	38189,92
9/25/2020	41982	38388,41
9/26/2020	42251	38581,15
9/27/2020	42448	38800,15
9/28/2020	42728	39047,69
9/29/2020	43003	39092,63
9/30/2020	43534	39911,2
10/1/2020	44135	40407,04
10/2/2020	44404	40346,36
10/3/2020	44723	40654,82
10/4/2020	44961	40740,49

Lampiran E. Hasil Akurasi

Kernel (70-30)	Radial	Linear	Sigmoid	Polynomial
Mape Train	150,551	22,467	2507,807	140,900
Mape Test	56,935	6,742	289,740	59,709
Cost Parameter	5	0,1	0,1	1