

Penerapan *Simple Linear Regression* Untuk Memprediksi Jumlah Kasus Covid Di Indonesia

Agus Kurniawan¹, Kurniawan Kokanda², Winderson^{3*}

^{1,2,3} Teknik Perangkat Lunak, Universitas Universal

* Corresponding author E-mail: windlim79@uvers.ac.id

Article Info

Article history:

Received 29-11-2021

Revised 07-12-2021

Accepted 08-12-2021

Keyword:

COVID-19,

Indonesia,

RStudio,

Simple linear regression.

ABSTRACT

The rapid spread of covid has had many impacts on all people in the world. In Indonesia, the first case of COVID-19 was found in March 2020 and continues to spread until now. Many government efforts started from PSBB to PPKM and also several regulations in it to prevent the spread of this virus. So, by using the RStudio program and trying to execute the spread of COVID-19 data with this simple linear regression method to predict the number of COVID-19 cases to find out whether there will be a case increase in Indonesia on a certain date. Besides that, the result of the predicted data can be used for any purpose or it can also use to compare predictive data with actual data.

Copyright © 2021 Journal of Digital Ecosystem for Natural Sustainability.
All rights reserved.

I. PENDAHULUAN

Coronavirus adalah salah satu keluarga virus penyebab penyakit yang menular pada hewan dan manusia. Pada manusia, beberapa mutasi Coronavirus ini diketahui menyebabkan infeksi pada saluran pernapasan. Gejala yang disebabkan oleh virus ini dimulai dari flu biasa, hingga penyakit yang lebih parah seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS). Penularan virus ini berlangsung antar manusia (human to human transmission) yang menimbulkan peningkatan jumlah kasus yang luar biasa hingga pada akhir Januari 2020 didapatkan peningkatan 2000 kasus terkonfirmasi dalam 24 jam. Pada akhir Januari 2020 World Health Organization (WHO) menetapkan status Global Emergency pada kasus virus Corona ini dan pada 11 Februari 2020 WHO menamakan virus ini sebagai COVID-19 [1], [2].

Penularan virus COVID-19 berasal dari orang yang sudah terinfeksi virus kepada orang lain di sekitarnya melalui percikan batuk atau bersin. COVID-19 juga dapat menular melalui benda-benda yang sudah terkontaminasi oleh percikan batuk maupun bersin dari penderita COVID-19 itu sendiri. Oleh karena itu, jika orang lain menyentuh benda-benda yang sudah terkontaminasi virus tersebut dan kemudian menyentuh area mata, hidung dan mulut maka mereka dapat tertular virus COVID-19 [3].

Virus penyebab COVID-19 dapat bertahan di udara sekitar satu jam, sedangkan di permukaan benda-benda dapat bertahan hingga beberapa jam. Di permukaan benda yang berbahan plastik dan besi tahan karat, virus ini dapat bertahan hingga 72 jam. Pada cardboard virus ini juga dapat bertahan selama 24 jam dan pada tembaga virus ini dapat bertahan hingga 4 jam [3].

Dengan kemampuan bertahan dan penyebaran yang cepat ini, penularan virus COVID-19 menjadi sangat cepat dan memberikan dampak yang sangat luas bagi seluruh masyarakat dari negara di seluruh dunia termasuk Indonesia. Di Indonesia kasus COVID-19 muncul pertama kali pada bulan Maret 2020. Sejak saat itu persebaran COVID-19 makin meluas hingga sekarang. Saat ini jumlah angka positif COVID-19 di Indonesia mencapai 4,07 juta jiwa dengan jumlah korban meninggal mencapai 132 ribu. Indonesia juga berjuang melawan COVID-19 dengan memodifikasi dan memperketat kebijakan mengenai karantina wilayah (lockdown) menjadi pembatasan sosial berskala besar (PSBB) yang bersifat lokal di seluruh kota Indonesia sesuai tingkat keparahan kasus di wilayah provinsi, kabupaten, atau kota [1], [4].

Kebijakan PSBB ini membawa dampak semakin terpuruknya kondisi ekonomi dan hal lainnya yang bersifat multidimensi. Menyadari kebijakan tersebut tidak dapat mengatasi pandemi COVID-19, mulai pertengahan Februari 2021 pemerintah melaksanakan kebijakan baru yang disebut dengan Pemberlakuan Pembatasan Kegiatan Masyarakat

(PPKM) berskala mikro serta mengupayakan vaksin massal kepada seluruh warga masyarakat Indonesia secara bertahap. Kegiatan PPKM dan vaksin kepada seluruh masyarakat ini bertujuan untuk [3], [5]:

- A. Mengatasi penyebaran COVID-19 yang sampai saat ini masih meningkat, pada tingkat desa dan kelurahan hingga kota.
- B. Mengoptimalkan pembatasan kegiatan masyarakat dalam berbagai aspek dengan pemberlakuan PPKM berskala mikro.
- C. Meningkatkan kesadaran masyarakat untuk melaksanakan dan menaati peraturan yang sudah dibuat berupa protokol kesehatan dengan menggunakan masker, mencuci tangan, *social distancing* serta PPKM berskala mikro.

Berdasarkan ulasan yang telah dijelaskan, adapun tujuan dari penelitian ini adalah memprediksi jumlah kasus harian, tingkat kesembuhan, dan angka kematian yang disebabkan oleh virus COVID-19 ini dalam jangka waktu tertentu, sehingga dapat digunakan untuk pengambilan keputusan baik masyarakat maupun pemerintah dalam menetapkan dan menjalankan kebijakan PPKM berskala mikro di Indonesia.

II. METODE

Penelitian ini akan diaplikasikan dengan bantuan program RStudio dan langkah pengerjaannya dibagi menjadi beberapa tahap yaitu pertama pengumpulan dan inisiasi data, memplot data dan mengecek korelasi variabel data, memasukkan fungsi prediksi model regresi linear, dan yang terakhir akan memplot hasil prediksi dari data yang sudah dieksekusi. Detail alur pelaksanaan penelitian ini adalah:

a. Pengumpulan Data

Data yang akan diolah dalam penelitian ini adalah data rekap COVID-19 di Indonesia dari tanggal 2 Maret 2020 sampai dengan 11 November 2021. Data ini diambil dari Kawal Covid-19, Kementerian Riset dan Teknologi - Badan Riset dan Inovasi Nasional Republik Indonesia (<https://sinta.ristekbrin.go.id/covid/datasets>) [6].

b. Deskripsi Data

Data yang telah diambil dari Kawal Covid-19, Kementerian Riset dan Teknologi - Badan Riset dan Inovasi Nasional Republik Indonesia pada rentang waktu 2 Maret 2020 sampai dengan 11 November 2021 menggunakan format excell dan memiliki 4 atribut yaitu:

Tabel 1. Deskripsi Atribut

Atribut	Deskripsi
Tanggal	Jumlah atau banyaknya kasus pertanggalnya
Kasus Harian	Banyaknya penambahan kasus positif yang terpapar COVID-19
Sembuh	Banyaknya orang yang sembuh dari COVID-19
Meninggal Dunia	Banyaknya korban yang meninggal dunia karena COVID-19

c. Eksekusi Data

Dataset akan dieksekusi dengan program RStudio yang nantinya akan menghasilkan plot dan data yang dibutuhkan untuk tahap berikutnya. Berikut ini adalah *Listing Program* untuk prediksi kasus harian, tingkat kesembuhan dan angka kematian jika dituliskan dalam program RStudio:

Program Prediksi Kasus Harian

```
#plot kasus harian
plot(dataset_rstudio$`kasus harian`,xaxt="n",
ylab="kasus harian", xlab="")
axis(1, labels=paste(dataset_rstudio$`tanggal
gabung`,sep=""), at=1:620, las=3)

cor(dataset_rstudio$`tanggal
gabung`,dataset_rstudio$`kasus harian`)

fit<-lm(dataset_rstudio$`kasus harian` ~
dataset_rstudio$`tanggal gabung`)

tanggalbaru<-
data.frame(dataset_databaru$`tanggal baru`)
khbaru<-predict(fit, newdata=tanggalbaru)
style<-c(rep(1,620), rep(2,620))

plot(c(dataset_rstudio$`kasus harian`, khbaru),
xaxt="n", ylab="kasus harian", xlab="", pch=style,
col=style)
axis(1, at=1:1240, las=3,
labels=c(paste(dataset_rstudio$`tanggal
gabung`,sep=""), dataset_databaru$`tanggal
baru`))

#split data latih dan uji
datalatih<-dataset_rstudio[1:310, 1:2]
datauji<-dataset_rstudio[311:620, 1:2]

fit<-lm(datalatih$`kasus harian`
~datalatih$`tanggal gabung`)

prediksi<-predict(fit, datauji[,1])
selisih<-abs(prediksi-datauji[,2])

MAD<-sum(selisih)/310
MSE<-sum(selisih^2)/310
MAPE<-sum((selisih/datauji[,2])*100)/310
```

Program Prediksi Tingkat Kesembuhan

```
#plot tingkat kesembuhan
plot(dataset_rstudio$`sembuh`,xaxt="n",
ylab="sembuh", xlab="")
axis(1, labels=paste(dataset_rstudio$`tanggal
gabung`,sep=""), at=1:620, las=3)

cor(dataset_rstudio$`tanggal
gabung`,dataset_rstudio$`sembuh`)

fit<-lm(dataset_rstudio$`sembuh` ~
dataset_rstudio$`tanggal gabung`)

sembuhbaru<-data.frame(dataset_databaru$`tanggal
baru`)
smbbaru<-predict(fit, newdata=sembuhbaru)
style<-c(rep(1,620), rep(2,620))

plot(c(dataset_rstudio$`sembuh`, smbbaru),
xaxt="n", ylab="sembuh", xlab="", pch=style,
col=style)
axis(1, at=1:1240, las=3,
labels=c(paste(dataset_rstudio$`tanggal
gabung`,sep=""), dataset_databaru$`tanggal
baru`))
```

```
#split data latih dan uji
datalatih<-dataset_rstudio[1:310, 1:4]
datauji<-dataset_rstudio[311:620, 1:4]

fit<-lm(datalatih$`Sembuh` ~datalatih$`tanggal
gabung` )

prediksi<-predict(fit, datauji[,1])
selisih<-abs(prediksi-datauji[,4])

MAD<-sum(selisih)/310
MSE<-sum(selisih^2)/310
MAPE<-sum((selisih/datauji[,2])*100)/310
```

Program Prediksi Angka Kematian

```
#plot angka kematian
plot(dataset_rstudio$`Sembuh`,xaxt="n",
plot(dataset_rstudio$`Meninggal Dunia`,xaxt="n",
ylab="Meninggal Dunia", xlab="")
axis(1, labels=paste(dataset_rstudio$`tanggal
gabung`,sep=""), at=1:620, las=3)
```

```
cor(dataset_rstudio$`tanggal
gabung`,dataset_rstudio$`Meninggal Dunia`)
fit <- lm(dataset_rstudio$`Meninggal Dunia` ~
dataset_rstudio$`tanggal gabung`)
```

```
angkadf<-data.frame(dataset_databaru$`tanggal
baru`)
angkabarur<-predict(fit, newdata=angkadf)
style<-c(rep(1,620), rep(2,620))
```

```
plot(c(dataset_rstudio$`Meninggal Dunia`,
angkabarur), xaxt="n", ylab="Meninggal Dunia",
xlab="", pch=style, col=style)
axis(1, at=1:1240, las=3,
labels=c(paste(dataset_rstudio$`tanggal
gabung`,sep=""), dataset_databaru$`tanggal
baru`))
```

```
#split data latih dan uji
datalatih<-dataset_rstudio[1:310, 1:6]
datauji<-dataset_rstudio[311:620, 1:6]
```

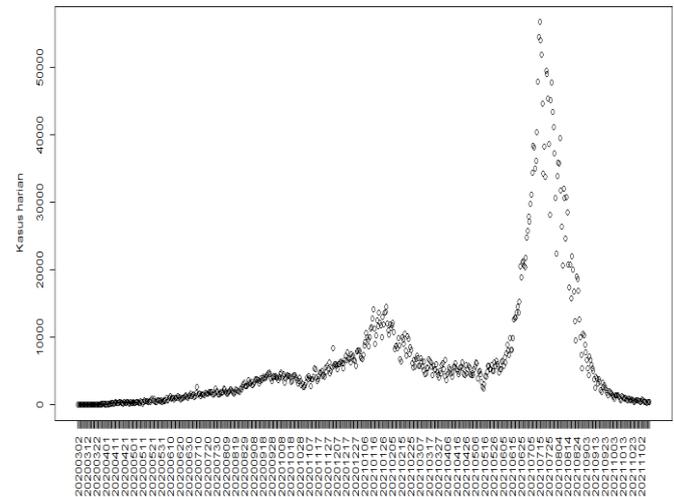
```
fit<-lm(datalatih$`Meninggal Dunia` ~
datalatih$`tanggal gabung` )
```

```
prediksi<-predict(fit, datauji[,1])
selisih<-abs(prediksi-datauji[,6])
```

```
MAD<-sum(selisih)/310
MSE<-sum(selisih^2)/310
MAPE<-sum((selisih/datauji[,2])*100)/310
```

III. HASIL DAN PEMBAHASAN

• Prediksi Kasus Harian



Gambar 1. Plot Actual Kasus Harian

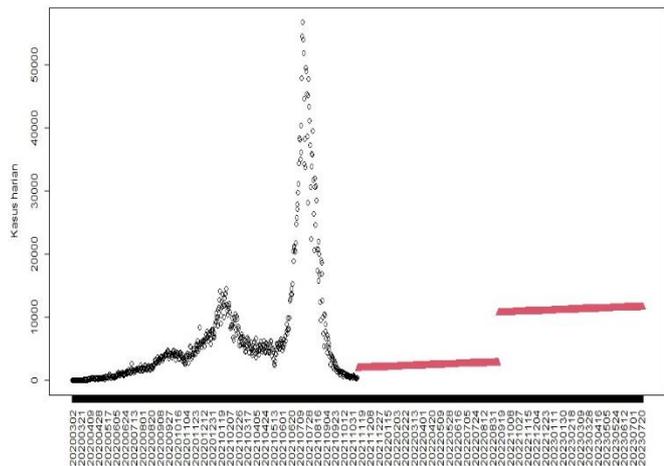
Gambar 1 merupakan visualisasi untuk data kasus harian covid di Indonesia yang dimulai dari 2 Maret 2021 sampai dengan 11 November 2021. Berikutnya hasil korelasi, MAD, MSE, dan MAPE untuk data kasus harian:

1. Korelasi = 0,4574
2. MAD = 9496,3840
3. MSE = 222460860
4. MAPE = 133,0013%

Untuk data kasus harian covid memiliki korelasi sebesar 0,4574, dimana jika korelasi mendekati nilai 0 maka variabel tanggal (x) dan angka kasus harian (y) mempunyai tingkat pengaruh yang rendah / hubungan antara kedua variabel tanggal (x) dan angka kasus harian (y) itu lemah ataupun tidak erat. Untuk kasus ini jika kita melihat nilai MAPE sebesar 133,0013% yang menunjukkan nilai kesalahan atau *error* yang cukup tinggi. Nilai MAPE yang tinggi ini dipengaruhi oleh variabel tanggal (x) yang memiliki angka kasus harian yang tinggi dan lonjakan yang besar dengan total kasus harian mencapai 56.757 kasus pada 15 Juli 2021.

Untuk data kasus harian ini memiliki *summary* seperti berikut:

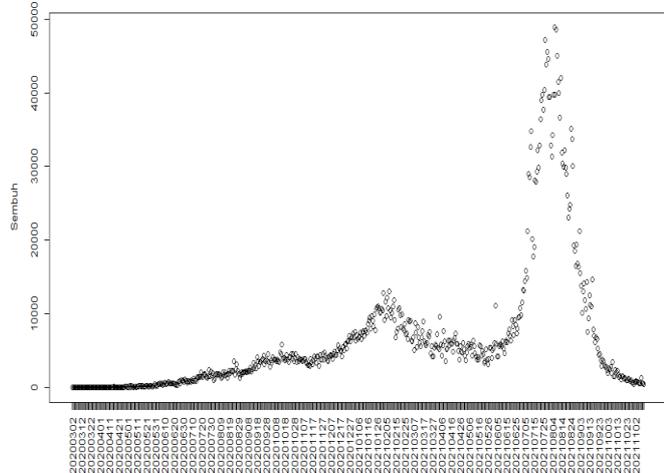
1. Kasus harian terendah adalah 0 kasus pada tanggal 3 Maret 2020, 4 Maret 2020, 5 Maret 2020, 7 Maret 2020, dan 12 Maret 2020.
2. Kasus harian tertinggi adalah 56.757 kasus yang terdapat pada tanggal 15 Juli 2021.
3. Untuk rata-rata kasus harian selama 620 hari adalah 6.854 kasus dengan total kasus hariannya mencapai 4.249.758 kasus.



Gambar 2. Plot *forecast* Kasus Harian

Berikutnya pada gambar 2 menunjukkan hasil visualisasi prediksi untuk data kasus harian yang dimulai dari 12 November sampai dengan 24 Juli 2023. Dimana perbandingan antara data aktual dan data prediksi memiliki perbandingan satu banding satu dengan jumlah data yang ada yaitu 620 tanggal. Jika kita lihat pergerakan atau penambahan dari prediksinya, data baru atau data hasil prediksinya bergerak dengan landai hingga 13 September 2022. Kemudian diprediksi naik hingga lima kali lipat dan bergerak kembali dengan landai hingga 24 Juli 2023.

• **Prediksi Angka Kesembuhan**



Gambar 3. Plot *Actual* Angka Kesembuhan

Dari Gambar ke-3 yang merupakan visualisasi plot data angka kesembuhan dari covid di Indonesia yang dimulai dari 2 Maret 2021 hingga 11 November 2021. Berikutnya hasil korelasi, MAD, MSE, dan MAPE untuk data angka kesembuhan :

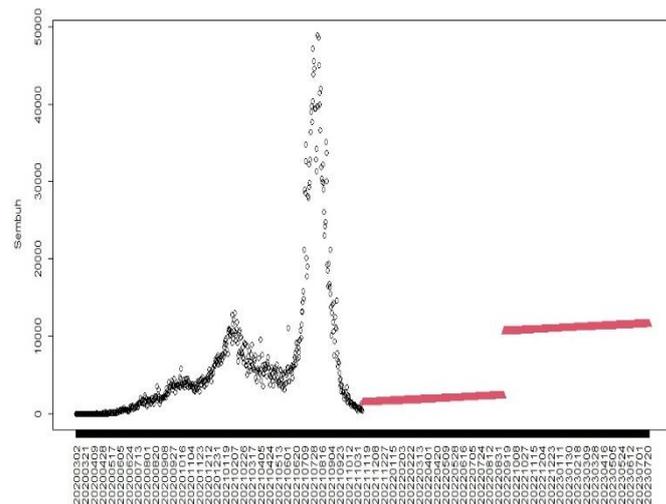
1. Korelasi = 0,5058
2. MAD = 9553,111
3. MSE = 206744583
4. MAPE = 131,7745%

Untuk data angka kesembuhan covid memiliki korelasi sebesar 0,5058, dimana sama seperti pada kasus harian covid

jika korelasi mendekati nilai angka 0 maka variabel tanggal (x) dan angka kesembuhan (y) mempunyai tingkat pengaruh yang rendah atau hubungan antara kedua variabel tanggal (x) dan angka kesembuhan (y) itu lemah ataupun tidak erat. Untuk kasus ini jika kita melihat nilai MAPE yang memiliki nilai sebesar 131,7745% itu menunjukkan nilai kesalahan atau *error* yang cukup tinggi. Nilai MAPE yang tinggi ini dipengaruhi oleh variabel tanggal (x) yang memiliki angka kesembuhan yang tinggi dan lonjakan yang besar pada tanggal 6 Agustus 2021 dengan angka kesembuhan mencapai 48.832 kasus.

Setelah mengetahui hasil korelasi, MAD, MSE, dan MAPE, berikut ini adalah hasil *summary* dari data kesembuhan:

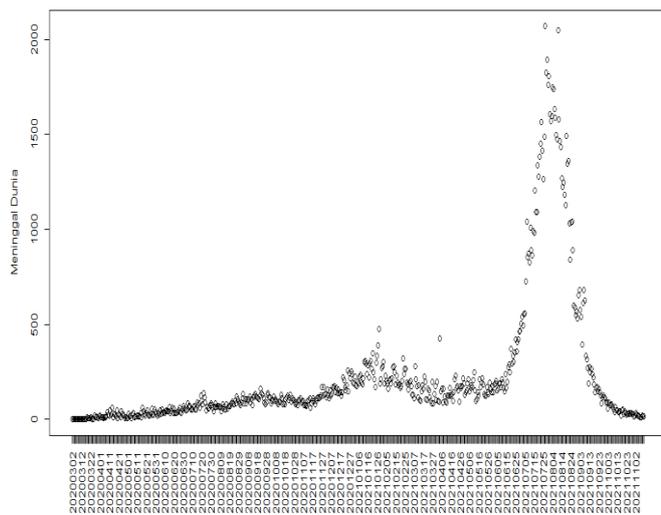
1. Tingkat kesembuhan terendah adalah 0 kasus sembuh pada tanggal 2 Maret 2020 sampai dengan 10 Maret 2020, 15 Maret 2020, 16 Maret 2020, dan 24 Maret 2020.
2. Untuk kasus sembuh tertinggi adalah 48.832 kasus dan terdapat pada tanggal 6 Agustus 2021.
3. Untuk rata-rata tingkat kasus kesembuhan selama 620 hari adalah 6.608 kasus dari total kasus kesembuhan mencapai 4.096.664 kasus sembuh.



Gambar 4. Plot *Forecast* Angka Kesembuhan

Pada Gambar 4 ini menunjukkan hasil visualisasi prediksi untuk data angka kesembuhan yang dimulai dari 12 November sampai dengan 24 Juli 2023. Dimana sama dengan perbandingan pada kasus harian sebelumnya, perbandingan antara data aktual dan data prediksi memiliki perbandingan satu banding satu dengan jumlah data yang ada yaitu 620 tanggal. Jika kita lihat pergerakan atau penambahan angka kesembuhan ini tidak jauh berbeda dari kasus harian dimana prediksinya bergerak dengan landai dari 12 November 2021 hingga 13 September 2022 diprediksi naik hingga lima kali lipat dan kembali bergerak dengan landai hingga 24 Juli 2023.

• Prediksi Angka Kematian



Gambar 5. Plot Actual Angka Kematian

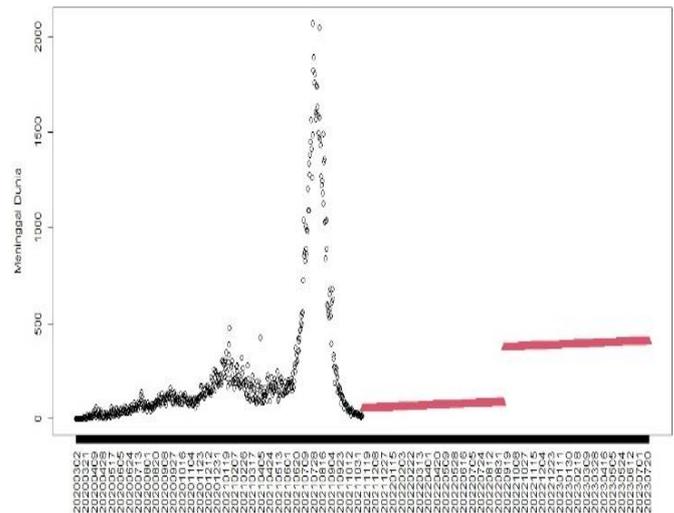
Kemudian untuk Gambar ke-5 ini yang merupakan visualisasi plot data angka kematian dari covid di Indonesia yang dimulai dari 2 Maret 2021 sampai dengan 11 November 2021. Berikutnya hasil korelasi, MAD, MSE, dan MAPE untuk data angka kematian :

1. Korelasi = 0,4413
2. MAD = 331,4596
3. MSE = 312670,5000
4. MAPE = 105,5577%

Untuk data angka kematian covid ini memiliki nilai korelasi sebesar 0,4413, dimana jika nilai korelasi yang mendekati nilai 0 maka variabel tanggal (x) dan angka kematian (y) mempunyai tingkat pengaruh yang rendah atau hubungan antara kedua variabel tanggal (x) dan angka kematian (y) itu lemah ataupun tidak erat. Untuk kasus ini jika kita melihat nilai MAPE sebesar 105,5577% yang merujuk terhadap nilai kesalahan atau *error* yang cukup tinggi. Nilai MAPE yang tinggi ini dipengaruhi oleh variabel tanggal (x) yang memiliki angka kematian cukup tinggi dan lonjakan besar dimana terjadi pada tanggal 27 Juli 2021 yang mencapai 2.069 korban jiwa.

Sebelum melihat data hasil prediksi, berikut adalah hasil *summary* untuk data angka kematian:

1. Angka kematian terendah terjadi pada tanggal 2 Maret 2020 sampai dengan 10 Maret 2020, 13 Maret 2020, 15 Maret 2020, dan 16 Maret 2020 dengan 0 kasus kematian.
2. Angka kematian tertinggi yaitu 2.069 kasus kematian terjadi pada tanggal 27 Juli 2021.
3. Untuk rata-rata angka kematian selama 620 hari adalah 232 kasus dari total angka kematian yang mencapai 143.608 kasus.



Gambar 6. Plot Forecast Angka Kematian

Untuk gambar terakhir, Gambar 6 yang merupakan plot terakhir hasil visualisasi prediksi untuk data angka meninggal yang dimulai dari 12 November sampai dengan 24 Juli 2023. Dimana perbandingan antara data aktual dan data prediksi memiliki perbandingan satu banding satu dengan jumlah data yang ada yaitu 620 tanggal. Jika kita lihat pergerakan atau penambahan bergerak dengan landai atau tidak signifikan hingga 13 September 2022, tetapi kemudian diprediksi naik hingga lima kali lipat dan kemudian kembali bergerak dengan landai hingga 24 Juli 2023.

• Sample Hasil Prediksi Kasus Covid

Untuk mendapatkan informasi penting dari hasil prediksi yang telah dilakukan, kita akan mengambil beberapa *sample* data dari hasil prediksi untuk tanggal penting atau tanggal merah yang dimulai dari Desember 2021 – April 2022 :

Tabel 2. Sample Hasil Prediksi

Tanggal	Hasil Prediksi		
	Kasus harian	Angka Kesembuhan	Angka Kematian
25 Desember 2021	2064	1589	57
1 Januari 2022	2070	1596	57
1 Februari 2022	2161	1691	60
28 Februari 2022	2247	1781	63
3 Maret 2022	2249	1784	64
15 April 2022	2413	1954	70

IV. KESIMPULAN

Melalui uji coba prediksi jumlah kasus covid di Indonesia dengan program Rstudio ini kita dapat menyimpulkan bahwa :

1. Data hasil prediksi baik untuk kasus harian, angka kesembuhan, dan angka kematian memiliki angka korelasi yang mendekati 0. Dimana ini berarti variabel tanggal (x) dan variabel kasus harian, angka kesembuhan, dan angka

kematian (y) memiliki hubungan yang lemah atau tidak erat. Dapat dilihat juga dari MAPE untuk ketiga prediksi itu melebihi 100% dan berarti data hasil prediksi ini memiliki kemungkinan *error* atau kesalahan data yang memungkinkan perbedaan data hasil prediksi dan data sesungguhnya.

2. Jika melihat hasil prediksi pada *sample* data untuk tanggal penting / tanggal merah memiliki kenaikan yang tidak begitu signifikan. Tetapi alangkah baiknya aturan-aturan yang telah ditetapkan yaitu PPKM tetap dijalankan. Ini mencegah supaya penyebaran COVID-19 di Indonesia tetap terkontrol dan tidak selalu meningkat.

UCAPAN TERIMA KASIH

Puji syukur kehadiran Tuhan Yang Maha Esa atas limpahan rezeki dan rahmat-Nya sehingga kami dapat menyelesaikan penulisan karya tulis ini. Tidak lupa juga kami ucapkan terima kasih atas bimbingan serta arahan dan juga waktunya kepada bapak Akhmad Rezki Purnajaya, S.Kom, M.Kom dan ibu Steffi Adam, S.SI., M.MSI.

DAFTAR PUSTAKA

- [1] Y. R. B. T. P. D. melitus T. I. 2019. 89-94 Sitepu and J. N. Simanungkalit, "Jurnal Penelitian Perawat Profesional," *Y sitepu*, vol. 1, no. November, pp. 89-94, 2019, [Online]. Available: <http://jurnal.globalhealthsciencegroup.com/index.php/JPPP/article/download/83/65>. [Accessed: Nov. 20, 2021]
- [2] P. D. O. Davies, "Multi-drug resistant tuberculosis," *CPD Infect.*, vol. 3, no. 1, pp. 9-12, 2002.
- [3] M. K. Sari, "Sosialisasi tentang Pencegahan Covid-19 di Kalangan Siswa Sekolah Dasar di SD Minggiran 2 Kecamatan Papan Kabupaten Kediri," *J. Karya Abdi*, vol. 4, no. 1, pp. 80-83, 2020.
- [4] K. Perencanaan *et al.*, "Covid-19, New Normal, dan Perencanaan Pembangunan di Indonesia," *J. Perenc. Pembang. Indones. J. Dev. Plan.*, vol. 4, no. 2, pp. 240-252, 2020.
- [5] U. Napitu, Corry, and M. K. D. Matondang, "Sosialisasi pembatasan pelaksanaan kegiatan masyarakat (PPKM) mikro di kelurahan bah Kapul," *Community Dev. J.*, vol. 2, no. 2, pp. 232-241, 2021.
- [6] KawalCOVID19, "COVID-19 di Indonesia @kawalCovid19.", Kementerian Riset dan Teknologi - Badan Riset dan Inovasi Nasional Republik Indonesia, 2020, [Online]. Tersedia: <https://sinta.ristekbrin.go.id/covid/datasets>. [Accessed: Nov. 14, 2021].