

Implementasi Data Mining untuk Prediksi Penderita Covid-19 Menggunakan Model Pertumbuhan Eksponensial

Tresna Dewi Fadiini*, Yurika Permanasari, Didi Suhaedi

Prodi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*tresnadewi78@gmail.com, yurikakoe@gmail.com, dsuhaedi@gmail.com

Abstract. The number of *Covid-19* sufferers in Indonesia continues to increase, especially in the provinces of DKI Jakarta, Bali and West Java. This increase is a problem for the community as well as the government. Important information from data on *Covid-19* sufferers can be extracted to gain knowledge by using data mining. Predicting the number of *Covid-19* sufferers that will come is the first step in anticipating this case. Data mining in this study serves to process raw data and displays predictive information for *Covid-19* sufferers in the 9th month and the accuracy of the model. The technique used to gain knowledge is to apply the differential equation model, namely the exponential growth model by looking at the observation time interval as a continuous variable. It can be concluded that data mining can provide information that the prediction of the increase in *Covid-19* will continue to rise, data mining can also provide information that the exponential growth model can calculate the prediction of the increase in *Covid-19*.

Keywords: *Covid-19, Exponential Growth Model, Data Mining.*

Abstrak. Angka penderita *Covid-19* di Indonesia terus menerus mengalami kenaikan terutama di provinsi DKI Jakarta, Bali, dan Jawa Barat. Kenaikan ini menjadi masalah bagi masyarakat dan juga pemerintah. Informasi penting dari data penderita *Covid-19* dapat digali untuk mendapatkan knowledge dengan menggunakan data mining. Memprediksi angka penderita *Covid-19* yang akan datang menjadi langkah awal dalamantisipasi kasus ini. Data mining pada penelitian ini berfungsi untuk mengolah data mentah dan menampilkan informasi prediksi penderita *Covid-19* pada bulan ke-9 dan nilai akurasi model. Teknik yang digunakan untuk mendapatkan knowledge adalah dengan menerapkan model persamaan diferensial yaitu model pertumbuhan eksponensial dengan memandang interval waktu pengamatan sebagai variabel yang kontinu. Dapat disimpulkan bahwa data mining dapat memberikan informasi bahwa prediksi kenaikan *Covid-19* akan terus naik, data mining juga dapat memberikan informasi bahwa model pertumbuhan eksponensial dapat menghitung prediksi kenaikan *Covid-19*.

Kata Kunci: *Covid-19, Model Pertumbuhan Eksponensial, Data Mining.*

A. Pendahuluan

Dua bulan setelah China melaporkan kasus Covid-19, Presiden Joko Widodo didampingi Menteri Kesehatan mengonfirmasi kasus Covid-19 pertama di Indonesia di Istana Kepresidenan, Jakarta, pada tanggal 2 (dua) Maret (Putri, 2020). Pertumbuhan penderita Covid-19 di Indonesia terus meningkat setiap harinya, dimana hal ini menjadi suatu masalah yang meresahkan baik bagi pemerintah, maupun seluruh masyarakat Indonesia. Dilihat dari data bulan Agustus 2020 terdapat 174.796 kasus terkonfirmasi positif Covid-19 (Anwar, 2020) hingga akhir tahun 2020 total kasus terkonfirmasi terus meningkat hingga 743.198 penderita (K, 2020).

Kondisi ini menjadi masalah besar di Indonesia terutama di beberapa provinsi seperti DKI Jakarta, Jawa Barat, dan Bali, maka diperlukan antisipasi pemerintah dan masyarakat untuk pengendalian kasus Covid-19 di masa yang akan datang. Salah satunya bisa dimulai dengan memprediksi angka penderita Covid-19. Fenomena yang ada dapat dianalisis dari berbagai sudut pandang, salah satunya dapat dilihat dalam bentuk model matematika. Contoh aplikasi matematika yang dapat diterapkan dalam kasus ini adalah pemodelan dengan persamaan diferensial, khususnya model pertumbuhan eksponensial (Kurniawan, Holisin, & Kristanti, 2017).

Model pertumbuhan eksponensial adalah model yang menggambarkan pola pertumbuhan populasi dengan memandang interval waktu pengamatan sebagai variabel yang kontinu (Rohaeni). Model pertumbuhan eksponensial dapat menunjukkan suatu informasi baru berupa prediksi angka penderita Covid-19. Model pertumbuhan eksponensial merupakan teknik analisis data yang dipilih oleh penulis karena data statistik kumulatif penderita Covid-19 di DKI Jakarta, Bali, dan Jawa Barat adalah data yang bersifat eksponensial (Aswicahyono, 2020).

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose, 2006). Salah satu fungsi dari data mining ialah menampilkan informasi baru (*knowledge*) yang bertujuan untuk memahami lebih jauh tentang data yang diamati. Adapun fungsi data mining dalam analisis data yaitu mengolah data mentah yang telah diperoleh melalui *pre-processing* untuk menghasilkan informasi baru. Dari beberapa fungsi tersebut data mining dapat membantu dalam proses analisis data untuk menghasilkan informasi tentang prediksi penderita Covid-19. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Untuk mengetahui *knowledge* dari data Covid-19.
2. Untuk mengetahui prediksi angka penderita Covid-19 pada bulan September 2021 di provinsi yaitu DKI Jakarta, Jawa Barat, dan Bali.

B. Metodologi Penelitian

Pada pengolahan data peneliti menggunakan metode data mining. Data yang digunakan adalah data penderita Covid-19 dari awal Januari 2021 hingga akhir Juni 2021 di 3 provinsi yakni DKI Jakarta, Jawa Barat, dan Bali.

Data-data yang digunakan diperoleh dari beberapa sumber yakni:

1. Data kumulatif harian penderita Covid-19 provinsi DKI Jakarta diperoleh dari situs <https://corona.jakarta.go.id/id/data-pemantauan>.
2. Data kumulatif harian penderita Covid-19 Terkonfirmasi Jawa Barat diperoleh dari situs <https://pikobar.jabarprov.go.id/>.
3. Data perkembangan akumulatif harian penderita Covid-19 provinsi Bali diperoleh dari situs <https://infocorona.baliprov.go.id/>.

Data yang telah diolah tadi diproses menggunakan model pertumbuhan eksponensial.

C. Hasil Penelitian dan Pembahasan Proses Mining



Gambar 1. Flowchart Process Mining

Data yang telah diolah sebelumnya melalui teknik transformasi data selanjutnya dimasukkan kedalam rumus untuk mencari nilai konstanta (k) dengan rumus yang telah disebutkan sebelumnya. Setelah mendapatkan nilai konstanta (k) selanjutnya dimasukkan kedalam rumus model pertumbuhan eksponensial untuk mendapatkan persamaan model pertumbuhan eksponensial yang akan digunakan untuk memprediksi penderit Covid-19.

Setelah didapatkan persamaan tersebut akan diuji galatnya dengan rumus MAPE. Skala persentase pada perhitungan MAPE menentukan apakah model ini bisa digunakan untuk memprediksi angka penderit Covid-19. Langkah terakhir yaitu menghitung prediksi penderit Covid-19 bulan ke-9 di provinsi DKI Jakarta, Bali, dan Jawa Barat.

Mean Absolute Percentage Error (MAPE)

Pada tahap ini menghitung nilai akurasi untuk model pertumbuhan eksponensial yang didapat dengan *Mean Absolute Percentage Error* (MAPE). Dalam menghitung MAPE dibutuhkan variabel nilai data historis pada periode ke- t , nilai prediksi pada periode ke- t dan banyaknya pengamatan. Dibawah ini merupakan perhitungan MAPE secara *manual*.

1. Menghitung MAPE untuk model pertumbuhan eksponensial provinsi DKI Jakarta menggunakan nilai aktual dan nilai prediksi bulan Mei dan bulan Juni.

$$MAPE = \frac{\sum_{t=1}^n \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n}$$

$$MAPE = \frac{\left| \left(\frac{430.059 - 472.712}{430.059} \right) 100 \right| + \left| \left(\frac{543.468 - 543.468}{543.468} \right) 100 \right|}{2}$$

$$MAPE = \frac{9,917941492 + 0}{2}$$

$$MAPE = 4,958970746$$

2. Menghitung MAPE untuk model pertumbuhan eksponensial provinsi Bali menggunakan nilai aktual dan nilai prediksi bulan Mei dan bulan Juni.

$$MAPE = \frac{\sum_{t=1}^n \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n}$$

$$MAPE = \frac{\left| \left(\frac{47.259 - 44.125}{47.259} \right) 100 \right| + \left| \left(\frac{50.217 - 50.217}{50.217} \right) 100 \right|}{2}$$

$$MAPE = \frac{6,631541082 + 0}{2}$$

$$MAPE = 3,315770541$$

3. Menghitung MAPE untuk model pertumbuhan eksponensial provinsi Jawa Barat menggunakan nilai aktual dan nilai prediksi bulan Mei dan bulan Juni.

$$MAPE = \frac{\sum_{t=1}^n \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n}$$

$$MAPE = \frac{\left| \left(\frac{313.949 - 316.541}{313.949} \right) 100 \right| + \left| \left(\frac{381.455 - 381.455}{381.455} \right) 100 \right|}{2}$$

$$MAPE = \frac{0,85746411 + 0}{2}$$

$$MAPE = 0,428732055$$

Dari perhitungan di atas didapat nilai MAPE seperti Tabel 1 berikut.

Tabel 1. Hasil perhitungan MAPE

Jakarta			
	Aktual	Prediksi	galat
Mei	430.059	47.279	9,917941492
Juni	543.468	543.468	0
		Mape	4,958970746
Bali			
Mei	47.259	44.125	6,631541082
Juni	50.217	50.217	0
		Mape	3,315770541
Jawa Barat			
Mei	313.949	316.641	0,85746411
Juni	381.455	381.455	0
		mape	0,428732055

sumber: perhitungan manual menggunakan Microsoft excel

Terdapat analisa tentang nilai MAPE sebagaimana tertulis dalam tabel di bawah ini.

Tabel 2. Range MAPE

Range MAPE	Arti nilai
<10%	Kemampuan model peramalan sangat baik
10-20%	Kemampuan model peramalan baik
20-50%	Kemampuan model peramalan layak
>50%	Kemampuan model peramalan buruk

Model Pertumbuhan Eksponensial

Untuk memprediksi angka penderita pada bulan ke-9 perlu dibuat model matematikanya terlebih

dahulu. Dari Rumus yang telah ada, diketahui yang diperlukan untuk memodelkan Prediksi Angka penderita Covid-19 di Indonesia ini yaitu P , t , P_0 , t_0 dengan dari data pada Tabel 3, maka diketahui sebagai berikut.

Tabel 3. komposisi analisis data

	Jakarta	Bali	Jawa Barat
P	543.468	50.217	381.455
P₀	269.718	26.304	150.336
t	6	6	6
t₀	1	1	1

sumber: diambil dari data yang telah diolah

Pada tabel 4 diketahui nilai P yang merupakan jumlah penderita pada bulan t , t , merupakan bulan ke-6, P_0 yang merupakan jumlah penderita pada bulan ke t_0 dan t_0 merupakan bulan ke-1. Dari data yang diketahui dihitung secara manual seperti dibawah ini.

1. Mencari nilai konstanta pembanding (k)
 - a. Nilai konstanta pembanding (k) untuk provinsi DKI Jakarta

$$k = \frac{1}{t - t_0} \left(\ln \frac{P}{P_0} \right)$$

$$k = \frac{1}{6 - 1} \left(\ln \frac{543.468}{269.718} \right)$$

$$k = 0,1401187717$$

- b. Nilai konstanta pembanding (k) untuk provinsi Bali

$$k = \frac{1}{t - t_0} \left(\ln \frac{P}{P_0} \right)$$

$$k = \frac{1}{6 - 1} \left(\ln \frac{50.217}{26.304} \right)$$

$$k = 0,993265192$$

- c. Nilai konstanta pembanding (k) untuk provinsi Jawa Barat

$$k = \frac{1}{t - t_0} \left(\ln \frac{P}{P_0} \right)$$

$$k = \frac{1}{6 - 1} \left(\ln \frac{381.455}{150.336} \right)$$

$$k = 0,1862240199$$

Setelah diperoleh nilai konstanta pembanding (k), nilai tersebut disubstitusikan kedalam model pertumbuhan eksponensial seperti pada Tabel 4.

Tabel 4. Hasil nilai k dan persamaan model pertumbuhan eksponensial

	Jakarta	Bali	Jawa Barat
k	0,1401187717	0,993265192	0,1862240199
P(t)	$P_0 e^{0,1401187717(t-t_0)}$	$P_0 e^{0,993265192(t-t_0)}$	$P_0 e^{0,1862240199(t-t_0)}$

sumber: perhitungan manual menggunakan Microsoft excel

Setelah mendapatkan model matematika, maka dapat memprediksikan jumlah penderita pada waktu yang ingin di prediksi yaitu bulan ke-9. Karena mencari jumlah penderita pada bulan ke 9, maka t yang digunakan yaitu 9 dan t_0 dan P_0 sama seperti pada Tahap 1. Menggunakan model yang telah diperoleh pada tabel 5, selanjutnya dilakukan analisis data sebagai berikut

1. Analisis data untuk provinsi DKI Jakarta

- a. $P(5) = P_0 e^{0,1401187717(t-t_0)}$
 $P(5) = 269.718 * e^{0,1401187717(5-1)}$
 $P(5) = 472.412$ jiwa
 - b. $P(6) = P_0 e^{0,1401187717(t-t_0)}$

- $$P(6) = 269.718 * e^{0,1401187717(6-1)}$$
- $$P(6) = 543.468 \text{ jiwa}$$
- c. $P(7) = P_0 e^{0,1401187717(t-t_0)}$
 $P(7) = 269.718 * e^{0,1401187717(7-1)}$
 $P(7) = 625.211 \text{ jiwa}$
- d. $P(8) = P_0 e^{0,1401187717(t-t_0)}$
 $P(8) = 269.718 * e^{0,1401187717(8-1)}$
 $P(8) = 719.249 \text{ jiwa}$
- e. $P(9) = P_0 e^{0,1401187717(t-t_0)}$
 $P(9) = 269.718 * e^{0,1401187717(9-1)}$
 $P(9) = 827.432 \text{ jiwa}$
2. Analisis data untuk provinsi Bali
- a. $P(5) = P_0 e^{0,993265192(t-t_0)}$
 $P(5) = 26.304 * e^{0,993265192(5-1)}$
 $P(5) = 44.125 \text{ jiwa}$
- b. $P(6) = P_0 e^{0,993265192(t-t_0)}$
 $P(6) = 26.304 * e^{0,993265192(6-1)}$
 $P(6) = 50.217 \text{ jiwa}$
- c. $P(7) = P_0 e^{0,993265192(t-t_0)}$
 $P(7) = 26.304 * e^{0,993265192(7-1)}$
 $P(7) = 57.150 \text{ jiwa}$
- d. $P(8) = P_0 e^{0,993265192(t-t_0)}$
 $P(8) = 26.304 * e^{0,993265192(8-1)}$
 $P(8) = 65.040 \text{ jiwa}$
- e. $P(9) = P_0 e^{0,993265192(t-t_0)}$
 $P(9) = 26.304 * e^{0,993265192(9-1)}$
 $P(9) = 74.019 \text{ jiwa}$
3. Analisis data untuk provinsi Jawa Barat
- a. $P(5) = P_0 e^{0,1862240199(t-t_0)}$
 $P(5) = 150.336 * e^{0,1862240199(5-1)}$
 $P(5) = 316.641 \text{ jiwa}$
- b. $P(6) = P_0 e^{0,1862240199(t-t_0)}$
 $P(6) = 150.336 * e^{0,1862240199(6-1)}$
 $P(6) = 381.455 \text{ jiwa}$
- c. $P(7) = P_0 e^{0,1862240199(t-t_0)}$
 $P(7) = 150.336 * e^{0,1862240199(7-1)}$
 $P(7) = 459.535 \text{ jiwa}$
- d. $P(8) = P_0 e^{0,1862240199(t-t_0)}$
 $P(8) = 150.336 * e^{0,1862240199(8-1)}$
 $P(8) = 553.599 \text{ jiwa}$
- e. $P(9) = P_0 e^{0,1862240199(t-t_0)}$
 $P(9) = 150.336 * e^{0,1862240199(9-1)}$
 $P(9) = 666.916 \text{ jiwa}$

Dari analisis data diatas diperoleh tabel seperti pada tabel 5, menunjukkan bahwa prediksi penderita Covid-19 pada bulan ke-9 (September) di provinsi DKI Jakarta sebanyak 827.432 jiwa, di provinsi Bali sebanyak 74.019 jiwa, dan di provinsi Jawa Barat sebanyak 666.916 jiwa.

Tabel 5. hasil analisis data

	Jakarta (jiwa)	Bali (jiwa)	Jawa Barat (jiwa)
P(5)	472412	44.125	316.641
P(6)	543.468	50.217	381.455
P(7)	625.211	57.150	459.535
P(8)	719.249	65.040	553.599
P(9)	827.432	74.019	666.916

sumber: perhitungan manual menggunakan Microsoft excel

D. Kesimpulan

Data mining dapat membantu dalam memberikan informasi bahwa prediksi kenaikan Covid-19 akan terus naik dilihat berdasarkan analisis data yang dilakukan, data mining juga memberikan informasi bahwa model eksponensial dapat memprediksi kenaikan kasus Covid-19 karena persentasi galat pada perhitungan MAPE menunjukkan skala <10% yang artinya model pertumbuhan eksponensial sangat baik dalam fenomena ini.

Acknowledge

Penelitian ini dapat dilaksanakan dengan baik berkat bantuan dari berbagai pihak, untuk itu peneliti mengucapkan terimakasih kepada sumber data yang telah menyediakan data untuk penelitian ini.

Daftar Pustaka

- [1] Anwar, F. (2020, Agustus 31). *Detik Health*. (detikcom) Retrieved Maret 6, 2021, from <https://health.detik.com/berita-detikhealth/d-5153716/update-corona-indonesia-31-agustus-tambah-2743-total-kasus-174796>
- [2] Aswicahyono, H. (2020). *Pertumbuhan Eksponensial dan Mitigasi Covid-19: Hindari Framing Dikotomis. SIS Commentaries DMRU-003*.
- [3] K, N. S. (2020, Desember 31). *Detik Health*. (detikcom) Retrieved Maret 6, 2021, from <https://health.detik.com/berita-detikhealth/d-5316472/update-corona-di-indonesia-31-desember-2020-tambah-8074-total-positif-743198>
- [4] Kurniawan, A., Holisin, I., & Kristanti, F. (2017). APLIKASI PERSAMAAN DIFERENSIAL BIASA MODEL EKSPONENSIAL DAN LOGISTIK PADA PERTUMBUHAN PENDUDUK KOTA SURABAYA. *Journal of Mathematics Education, Science and Tchnology*, 2, 129-141.
- [5] Larose, D. T. (2006). *Data Mining Methods and Models*. Hoboken New Jersey: John Willey & Sons, Inc.
- [6] Putri, G. S. (2020, November 23). *Menelusuri Klaster Pertama Penularan Covid-19 di Indonesia*. (Kompascom) Retrieved Maret 8, 2021, from <https://www.kompas.com/sains/read/2020/10/23/090200623/menelusuri-klaster-pertama-penularan-Covid-19-di-indonesia?page=all>
- [7] Rohaeni, O. (n.d.). *Model pertumbuhan Populasi Satu Spesies dengan Tundaan Waktu Diskrit*. Retrieved Maret 2, 2021, from <https://ejournal.unisba.ac.id/index.php/matematika/article/download/2541/172>.
- [8] Shafana Ninda Rizkya, Gunawan Gani, (2021). *Analisis Faktor yang Mempengaruhi Kesadaran Penduduk dalam Vaksin Covid-19 Menggunakan Metode Multivariate Adaptive Regression Spline*. *Jurnal Riset Matematika*, 1(2), 154-162.