

SKRIPSI

ESTIMASI PARAMETER-PARAMETER TIDAK KONSTAN
MODEL SIRD DENGAN METODE SEGMENTASI DAN
ARTIFICIAL NEURAL NETWORK (ANN) UNTUK DATA
PENYEBARAN COVID-19 DI INDONESIA



ALDYNOVA SURYADI

NPM: 6161901058

PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2023

FINAL PROJECT

**ESTIMATION OF NON-CONSTANT PARAMETERS OF THE
SIRD MODEL WITH SEGMENTATION AND ARTIFICIAL
NEURAL NETWORK (ANN) METHODS FOR DATA ON THE
SPREAD OF COVID-19 IN INDONESIA**



ALDYNNOVA SURYADI

NPM: 6161901058

**DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2023**

LEMBAR PENGESAHAN

ESTIMASI PARAMETER-PARAMETER TIDAK KONSTAN MODEL SIRD DENGAN METODE SEGMENTASI DAN *ARTIFICIAL NEURAL NETWORK (ANN)* UNTUK DATA PENYEBARAN COVID-19 DI INDONESIA

ALDYNNOVA SURYADI

NPM: 6161901058

Bandung, 31 Juli 2023

Menyetujui,

Pembimbing 1

Maria Anestasia, M.Si., M.Act.Sc.

Pembimbing 2

Dr. Andreas Parama Wijaya

Ketua Penguji

Farah Kristiani, Ph.D.

Anggota Penguji

Robyn Irawan, M.Sc.

Mengetahui,

Ketua Program Studi

Dr. Livia Owen

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

ESTIMASI PARAMETER-PARAMETER TIDAK KONSTAN MODEL SIRD DENGAN METODE SEGMENTASI DAN *ARTIFICIAL NEURAL NETWORK (ANN)* UNTUK DATA PENYEBARAN COVID-19 DI INDONESIA

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
31 Juli 2023



Aldynova Suryadi
NPM: 6161901058

ABSTRAK

Penyebaran penyakit COVID-19 merupakan masalah global yang berdampak besar, sehingga masalah tersebut perlu diteliti lebih lanjut untuk menekan dampak jika terjadi pandemi se-rupa di masa mendatang. Penyebaran penyakit COVID-19 di Indonesia dapat dimodelkan dengan model SIRD. Model tersebut membagi manusia menjadi empat kompartemen: manusia rentan, terinfeksi, sembuh, dan meninggal. Informasi dan strategi untuk mengantisipasi penyebaran penyakit dapat disimpulkan jika parameter-parameter pada model SIRD ini diketahui. Oleh sebab itu, pada skripsi ini parameter-parameter model SIRD diestimasi dengan segmentasi dan pembelajaran mesin. Parameter-parameter model SIRD tersebut menginterpretasikan tingkat penyebaran penyakit, kesembuhan, dan kematian yang tidak selalu konstan sepanjang waktu, sehingga hasil estimasi parameter-parameter merupakan fungsi tidak konstan. Parameter-parameter model SIRD yang diestimasi dengan segmentasi dapat dianalisa lebih lanjut berdasarkan kejadian-kejadian di dunia nyata. Berdasarkan analisa tersebut terdapat relevansi vaksinasi, libur lebaran, dan masuknya varian COVID-19 Omicron terhadap perubahan parameter-parameter model SIRD. Parameter-parameter model SIRD yang diestimasi dengan segmentasi diaplikasikan untuk peramalan sebulan mendatang. Berdasarkan peramalan yang dilakukan, didapat MAPE total yang relatif rendah yaitu 6,9%.

Kata-kata kunci: model SIRD; estimasi parameter; peramalan; pembelajaran mesin, segmentasi.

ABSTRACT

The spread of COVID-19 disease is a global problem that has a big impact, so the problem needs to be studied further to reduce the impact if there is a similar pandemic in the future. The spread of COVID-19 disease in Indonesia can be modeled with the SIRD model. The model divides humans into four compartments: susceptible, infected, recovered, and deceased. Information and strategies to anticipate the spread of the disease can be concluded if the parameters in this SIRD model are known. Therefore, in this thesis, the parameters of the SIRD model are estimated by segmentation and machine learning. The parameters of the SIRD model interpret the rates of disease spread, cure, and death that are not always constant over time, so the result of parameters estimation are non-constant functions. The parameters of the SIRD model estimated by segmentation can be further analyzed based on real-world events. Based on this analysis, there is an effect of vaccination, Eid holidays, and the inclusion of the COVID-19 omicron variant on changes in the parameters of the SIRD model. The SIRD model parameters estimated by segmentation are applied for forecasting in the coming month. Based on the forecasting performed, a relatively low total MAPE of 6,9% is obtained.

Keywords: SIRD model; parameters estimation; forecasting; machine learning; segmentation.

DAFTAR ISI

KATA PENGANTAR	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 <i>State of the Art</i>	3
1.5 Batasan Masalah	3
1.6 Metodologi	3
1.7 Sistematika Pembahasan	4
2 LANDASAN TEORI	5
2.1 Sistem Persamaan Diferensial Biasa	5
2.2 Metode Runge-Kutta	6
2.3 Model SIRD	7
2.4 Hubungan Antar Kompartemen Model SIRD	9
2.5 Analisa Titik-Titik Perubahan	10
2.6 <i>Trust Region Interior Reflective Method</i>	10
2.6.1 Fungsi Kuadratik	11
2.6.2 Ukuran <i>Trust Region</i>	11
2.7 Norm dari Matriks	12
2.8 Metrik Eror	12
2.9 <i>Artificial Neural Network</i>	13
2.9.1 Mekanisme Perhitungan ANN	14
2.9.2 Fungsi Aktivasi	15
2.9.3 Fungsi Biaya	16
2.9.4 <i>Optimizer</i>	16
2.9.5 <i>Hyperparameter ANN</i>	17
3 ESTIMASI PARAMETER-PARAMETER MODEL SIRD DARI DATA SINTETIS	18
3.1 Pembangkitan Data Sintetis	19
3.2 Estimasi Parameter-Parameter Model SIRD dengan Segmentasi	21
3.2.1 Estimasi Titik-Titik Perubahan	21
3.2.2 Hasil Estimasi Parameter-Parameter Model SIRD dengan Segmentasi	23
3.3 Estimasi Parameter-Parameter Model SIRD dengan Pembelajaran Mesin	25
3.3.1 Pembentukan Arsitektur ANN	25

3.3.2	Mekanisme Pembaruan Bobot dan Bias	26
3.3.3	<i>Hyperparameter ANN</i>	27
3.3.4	Hasil Estimasi dengan ANN	27
3.4	Perbandingan Estimasi dengan Segmentasi dan dengan Pembelajaran Mesin	30
4	ESTIMASI PARAMETER-PARAMETER MODEL SIRD BERDASARKAN DATA ASLI	32
4.1	Persiapan Data Asli	32
4.1.1	Pengolahan Data Asli	32
4.1.2	Pembagian Data Asli	33
4.2	Estimasi Parameter-Parameter Model SIRD dari Data Asli dengan Segmentasi	34
4.3	Perbandingan Hasil Simulasi Model SIRD dengan Data Asli	38
4.4	Estimasi Parameter-Parameter Model SIRD dari Data Asli dengan Pembelajaran Mesin	40
4.5	Peramalan Satu Bulan Mendatang Data Asli untuk Estimasi dengan Segmentasi	41
4.5.1	Parameter-Parameter Skenario Pertama	41
4.5.2	Parameter-Parameter Skenario Kedua	42
4.5.3	Parameter-Parameter Skenario Ketiga	43
4.5.4	Penentuan Parameter-Parameter Peramalan Terbaik	44
5	KESIMPULAN DAN SARAN	47
5.1	Kesimpulan	47
5.2	Saran	47
DAFTAR REFERENSI		49
A GAMBAR TAMBAHAN		51
B TABEL PENGUJIAN <i>Hyperparameter</i> TERBAIK		52

DAFTAR GAMBAR

2.1	Diagram transisi model SIRD.	7
2.2	Ilustrasi arsitektur ANN dengan 1 buah lapisan tersembunyi.	13
2.3	Fungsi aktivasi sigmoid.	15
2.4	Fungsi aktivasi tanh.	15
2.5	Fungsi aktivasi ReLU.	16
3.1	Diagram alir estimasi parameter-parameter model SIRD dari data sintetis.	18
3.2	Parameter-parameter model SIRD dari data sintetis.	20
3.3	Data sintetis dengan garis kuning: banyak manusia rentan, garis ungu: banyak manusia terinfeksi, garis hijau: banyak manusia sembuh, dan garis merah: banyak manusia meninggal.	20
3.4	Plot $\frac{R(t_i)}{N}$ dengan $\ln(S(t_i))$ dari data sintetis.	21
3.5	Nilai $grad_i$ untuk Gambar 3.4.	22
3.6	Gambar 3.4 yang disegmentasi berdasarkan waktu titik-titik perubahannya.	22
3.7	Parameter-parameter estimasi dan sebenarnya dari data sintetis dengan segmentasi.	24
3.8	Persentase eror absolut kompartemen manusia rentan, terinfeksi, sembuh, dan meninggal dari data sintetis dengan segmentasi.	25
3.9	Arsitektur ANN untuk mengestimasi parameter-parameter model SIRD.	26
3.10	Nilai MAPE tiap iterasi untuk garis biru: fungsi aktivasi tanh, garis jingga: fungsi aktivasi sigmoid, garis hijau: fungsi aktivasi ReLU.	26
3.11	Perbandingan parameter-parameter estimasi dan sebenarnya dari data asli dengan ANN.	28
3.12	Hasil prediksi dengan ANN dari data sintetis.	29
3.13	Persentase eror data prediksi dan data sintetis dengan ANN.	30
4.1	Data asli kompartemen manusia rentan, terinfeksi, sembuh, dan meninggal di Indonesia pada periode 24 Mei 2020 sampai 16 September 2022.	33
4.2	Data asli pada Gambar 4.1 yang dibagi menjadi data latih: titik berwarna biru, dan data uji: titik berwarna jingga.	34
4.3	Plot $\frac{R(t)}{N}$ dengan $\ln(S(t))$ dari data asli dengan segmentasi.	35
4.4	Parameter-parameter $\hat{\alpha}(t)$, $\hat{\beta}(t)$, dan $\hat{\gamma}(t)$ dari data asli.	37
4.5	Plot data asli dan data prediksi.	39
4.6	Persentase eror dari data asli dengan data prediksi.	40
4.7	Skenario pertama: parameter-parameter estimasi berwarna biru dan parameter-parameter untuk peramalan berwarna merah.	42
4.8	Skenario kedua: parameter-parameter estimasi berwarna biru dan parameter-parameter untuk peramalan berwarna merah.	43
4.9	Skenario ketiga: parameter-parameter estimasi berwarna biru dan parameter-parameter untuk peramalan berwarna merah.	44
4.10	Skenario kedua: perbandingan data uji berwarna biru dan hasil peramalan berwarna jingga.	45
4.11	Persentase eror absolut hasil peramalan tanggal 16 Agustus 2022 sampai 16 September 2022.	46

- A.1 Parameter-parameter estimasi dan banyak manusia terinfeksi yang disegmentasi berdasarkan kejadian-kejadian di dunia nyata, yaitu segmen berwarna hijau: vaksinasi periode pertama, segmen berwarna merah: libur lebaran dan pasca lebaran, segmen berwarna biru: vaksinasi periode kedua, dan segmen berwarna kuning: masuknya varian Omicron ke Indonesia.

51



DAFTAR TABEL

4.1 Pengujian nilai penalti dan banyak data minimal tiap segmen.	35
4.2 Nilai parameter-parameter tiap segmennya.	36
4.3 Penentuan <i>hyperparameter</i> estimasi parameter-parameter model SIRD dengan ANN.	41
4.4 Penentuan parameter-parameter terbaik untuk peramalan pada tanggal 16 Agustus 2022 sampai 16 September 2022.	45
B.1 Pengujian <i>hyperparameter</i> ANN untuk estimasi parameter-parameter model SIRD dari data sintetis dengan ANN.	52

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Black swan event adalah kejadian yang kemungkinan terjadinya sangat rendah namun memberikan dampak yang besar dalam kehidupan manusia dengan tiga ciri utama: sulit untuk diprediksi, memberikan dampak yang besar, dan setelah kejadian terjadi akan dicari polanya [1]. Baru-baru ini dunia mengalami *black swan event*, yakni penyebaran penyakit dengan nama *Coronavirus disease* 2019 atau disebut COVID-19. Kasus pertama dari COVID-19 ditemukan di Wuhan, ibu kota dari provinsi Hubei, Cina. Wabah ini kemudian ditetapkan sebagai pandemi global oleh *World Health Organization* (WHO) pada 11 Maret 2020 saat jumlah kasus mencapai 118.000 kasus dan telah menyebar ke 110 negara lainnya [2]. Penyakit COVID-19 dapat ditularkan melalui air liur yang terciprat saat batuk, bersin, dan berbicara. Karena penularannya yang begitu mudah, banyaknya kasus yang tercatat sampai 25 Oktober 2022 adalah 633.252.760 kasus, dengan 1,04% dari kasus yang tercatat meninggal dunia. Pemerintah dari setiap negara di seluruh dunia bahkan sudah melakukan aksi untuk membatasi kontak antara manusia, sehingga penyebaran COVID-19 dapat menurun.

Indonesia merupakan salah satu negara yang terdampak COVID-19, dengan kasus pertama terjadi pada tanggal 2 Maret 2020. Banyaknya kasus penderita COVID-19 yang tercatat sampai tanggal 22 Oktober 2022 adalah 6.467.189 dengan banyaknya penderita yang telah sembuh adalah sebanyak 6.289.633 dan yang telah meninggal karena COVID-19 sebanyak 158.398. Berdasarkan data tersebut tingkat kematian akibat COVID-19 berkisar 2,5%. Pemerintah Indonesia telah melakukan banyak tindakan untuk menurunkan tingkat penyebaran COVID-19 seperti pembatasan sosial berskala besar, himbauan untuk menjalankan protokol kesehatan, dan vaksinasi. Meskipun demikian, perlu diantisipasi terjadinya penyebaran penyakit serupa di masa mendatang, sehingga tingkat kematianya dapat diturunkan karena sudah dilakukan penanganan sebelum terjadi pandemi.

Salah satu cara mengantisipasi penyakit serupa di masa mendatang adalah dengan memanfaatkan pemodelan matematis untuk memprediksi perilaku penyebaran COVID-19, sehingga dapat ditentukan penanganan yang paling efektif. Penyebaran penyakit menular dapat dimodelkan dengan model matematis epidemiologi. Salah satunya adalah model kompartemen SIR, yaitu model yang membagi populasi dalam beberapa kompartemen tertentu dan menggambarkan hubungan dinamis antar kompartemennya. Model SIR ini membagi populasi manusia ke dalam tiga kompartemen berbeda, yaitu *S* sebagai kompartemen manusia rentan (*susceptible*), *I* sebagai kompartemen manusia yang terinfeksi virus (*infected*), dan *R* sebagai kompartemen manusia yang sembuh (*recovered*) [3].

Model lain yang merupakan pengembangan dari model SIR adalah model SIRD, yang membagi populasi manusia ke dalam empat kompartemen berbeda, yaitu kompartemen pada model SIR dengan tambahan D sebagai kompartemen manusia yang meninggal (*deceased*). Kompartemen S pada model SIRD dapat bertransisi menjadi kompartemen I , sedangkan kompartemen I dapat bertransisi menjadi kompartemen R atau kompartemen D . Tingkat terjadinya transisi pada model SIRD ini dapat direpresentasikan oleh parameter-parameter yang menyatakan tingkat penyebaran penyakit, tingkat kesembuhan, dan tingkat kematian [2]. Informasi dan strategi untuk menganalisis penyebaran penyakit dapat disimpulkan jika parameter-parameter ini diketahui nilainya. Oleh sebab itu, penentuan parameter-parameter model SIRD dari suatu data penyebaran penyakit menjadi salah satu masalah yang perlu perhatian khusus. Parameter-parameter model SIRD ini tidak selalu bernilai konstan sepanjang waktu. Jika parameter-parameter tersebut dimodelkan dengan fungsi konstan bagian demi bagian, maka kita dapat menggunakan metode dengan prinsip segmentasi.

Metode pembelajaran mesin dapat digunakan di mana kita tidak perlu memodelkan bentuk fungsi parameter-parameter [4, 5, 6]. Salah satu metode pembelajaran mesin yang dapat digunakan adalah *artificial neural network* (ANN). Metode ini sudah cukup berkembang dalam penyelesaian suatu persamaan diferensial biasa, serta suatu sistem persamaan diferensial biasa [4, 7]. Lebih lanjut, karena metode ANN memiliki prinsip pencarian parameter bobot dan bias, maka metode ini dapat diterapkan dalam pencarian parameter-parameter dari suatu sistem persamaan diferensial jika diberikan data [4, 7, 8]. Hal ini dapat dikatakan sebagai proses kebalikan. Dengan perkembangan yang begitu pesat terhadap permasalahan persamaan diferensial, maka metode ANN dibahas pada skripsi ini.

Pada skripsi ini, parameter-parameter model SIRD terlebih dahulu diestimasi menggunakan data sintetis, yaitu data yang merupakan solusi numerik model SIRD dengan parameter-parameter dan nilai awal yang diketahui. Estimasi terhadap data sintetis ini bertujuan untuk menentukan keakuratan metode yang digunakan sebelum parameter-parameter model SIRD dari data asli diestimasi. Data asli yang digunakan untuk mengestimasi parameter-parameter model SIRD adalah data penyebaran COVID-19 di Indonesia. Parameter-parameter model SIRD yang sudah diestimasi dari data asli akan diaplikasikan untuk peramalan sebulan mendatang.

1.2 Rumusan Masalah

Masalah yang dapat dirumuskan dalam skripsi ini adalah:

1. Bagaimana cara mengestimasi parameter-parameter berbentuk fungsi konstan bagian demi bagian pada model SIRD dengan segmentasi?
2. Bagaimana cara mengestimasi parameter-parameter model SIRD dengan pembelajaran mesin ANN?
3. Bagaimana cara melakukan peramalan penyebaran COVID-19 dari data asli selama satu bulan mendatang?

1.3 Tujuan

Tujuan dari penulisan skripsi ini adalah:

1. Menjabarkan langkah-langkah serta pembuatan kode program untuk mengestimasi parameter-parameter model SIRD yang berbentuk fungsi konstan bagian demi bagian dengan segmentasi.
2. Menjabarkan langkah-langkah serta pembuatan kode program untuk mengestimasi parameter-parameter model SIRD dengan menggunakan pembelajaran mesin ANN.
3. Menjabarkan langkah-langkah serta pembuatan kode program untuk meramalkan penyebaran COVID-19 dari data asli selama satu bulan mendatang.

1.4 *State of the Art*

Skripsi ini membahas estimasi parameter-parameter model SIRD dari data sintetis, yang tidak dibahas pada [2] dan [7]. Tujuan dilakukan estimasi terhadap data sintetis ini adalah untuk menentukan keakuratan serta keberhasilan metode estimasi pada data yang mengikuti model SIRD dengan parameter-parameter dan nilai awal yang diketahui. Metode estimasi dengan segmentasi yang dibahas pada [2] dan [7] memiliki kekurangan, yaitu diperlukan asumsi parameter-parameter model SIRD yang berbentuk fungsi konstan bagian demi bagian. Oleh sebab itu, dikembangkan metode estimasi dengan pembelajaran mesin yang dibahas pada [4]. Metode estimasi dengan pembelajaran mesin memungkinkan untuk mengestimasi parameter-parameter tidak konstan tanpa perlu asumsi bentuk dari parameter-parameternya. Skripsi ini juga membahas mengenai beberapa skenario peramalan dari data asli yang sebelumnya tidak dibahas pada [2].

1.5 Batasan Masalah

Semua data yang digunakan pada skripsi ini diasumsikan dari populasi yang konstan. Selain itu, kompartemen manusia sembuh diasumsikan tidak dapat bertransisi kembali menjadi manusia rentan.

1.6 Metodologi

Penelitian dalam skripsi ini diawali dengan membangkitkan data sintetis yang merupakan solusi numerik model SIRD dengan parameter-parameter dan nilai awal yang sudah ditetapkan. Parameter-parameter model SIRD akan diestimasi dari data sintetis tersebut dengan segmentasi dan pembelajaran mesin. Hasil estimasi dari kedua metode tersebut akan dievaluasi keakuratannya dan akan dibandingkan kelebihan dan kekurangan dari kedua metode tersebut. Selanjutnya, metode estimasi dengan segmentasi dan pembelajaran mesin akan diaplikasikan untuk mengestimasi parameter-parameter model SIRD dari data asli yang sudah diolah. Hasil estimasi dari data asli ini akan dianalisa berdasarkan kejadian-kejadian yang terjadi di Indonesia. Penelitian ini diakhiri dengan peramalan penyebaran COVID-19 sebulan mendatang.

1.7 Sistematika Pembahasan

Sistematika pembahasan dalam skripsi ini adalah sebagai berikut:

1. Bab 2: Landasan Teori

Bab ini membahas mengenai teori-teori yang mendukung skripsi ini, yaitu sistem persamaan diferensial biasa, metode Runge-Kutta, model SIRD, analisa titik-titik perubahan, metode *trust region interior reflective*, norm dari matriks, metrik error, dan *artificial neural network*.

2. Bab 3: Estimasi Parameter-Parameter Model SIRD dari Data Sintetis

Bab ini membahas mengenai langkah-langkah dalam membangkitkan data sintetis, mengestimasi parameter-parameter model SIRD dengan segmentasi dan pembelajaran mesin dari data sintetis yang sudah dibangkitkan. Selain itu, dibahas juga mengenai kelebihan dan kekurangan dari kedua metode tersebut.

3. Bab 4: Estimasi Parameter-Parameter Model SIRD dari Data Asli

Bab ini membahas mengenai persiapan data asli, mengestimasi parameter-parameter model SIRD dengan segmentasi, dan dengan pembelajaran mesin dari data asli. Selain itu, dibahas juga mengenai peramalan sebulan mendatang dari data asli.

4. Bab 5: Kesimpulan dan Saran

Bab ini membahas mengenai kesimpulan dari skripsi ini, serta saran untuk pengembangan dari skripsi ini.