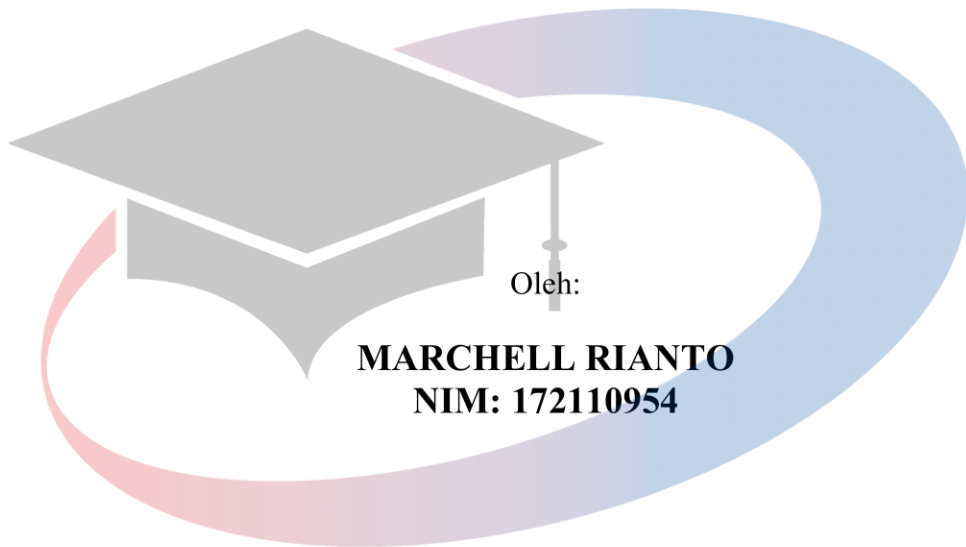


**ANALISIS RUNTUN WAKTU UNTUK MEMPREDIKSI
PANDEMI COVID-19 DENGAN MODEL RANDOM FOREST**

SKRIPSI

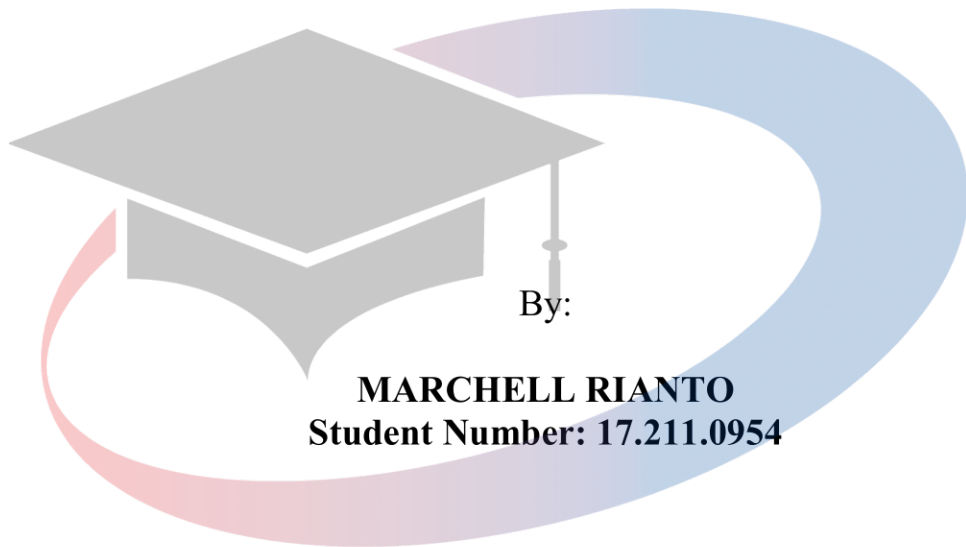


**UNIVERSITAS
MIKROSKIL**

**PROGRAM STUDI SISTEM INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2021**

**TIME-SERIES ANALYSIS FOR PREDICTING COVID-19
PANDEMIC WITH RANDOM FOREST MODEL**

FINAL RESEARCH



**UNIVERSITAS
MIKROSKIL**



**STUDY PROGRAM OF INFORMATION SYSTEM
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2021**

LEMBARAN PENGESAHAN

**ANALISIS RUNTUN WAKTU UNTUK MEMPREDIKSI
PANDEMI COVID-19 DENGAN MODEL RANDOM FOREST**

SKRIPSI

Diajukan Untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana Strata Satu
Program Studi Sistem Informasi

Oleh:

MARCHELL RIANTO
NIM: 172110954

Disetujui Oleh:

Dosen Pembimbing,

**UNIVERSITAS
MIKROSKIL**

Roni Yunis, S.Kom., M.T.

Medan, 1 Februari 2021

Diketahui dan Disahkan Oleh:

Ketua Program Studi
Sistem Informasi,

Rin Rin Meilani Salim, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Jurusan/Program Studi S-1 Sistem Informasi STMIK Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Marchell Rianto
NIM : 172110954
Peminatan : Sistem Informasi E-Bisnis

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir : Analisis Runtun Waktu Untuk Memprediksi Pandemi COVID-19 Dengan Model *Random Forest*
Tempat Penelitian : STMIK-STIE Mikroskil Medan
Alamat Tempat Penelitian : JL. M.H.Thamrin No.140, Kota Medan

Sehubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenarnya bahwa penelitian dan penulisan Tugas Akhir tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh STMIK Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada STMIK Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas hasil penelitian saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, STMIK Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan Tugas Akhir saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 5 Februari 2021

Saya yang membuat pernyataan,



Marchell Rianto

ANALISIS RUNTUN WAKTU UNTUK MEMPREDIKSI PANDEMI COVID-19 DENGAN MODEL RANDOM FOREST

Abstrak

Pandemi COVID-19 yang muncul pada awal tahun 2020 saat ini telah menyebar ke seluruh dunia dan mempengaruhi segala aspek kehidupan manusia. Di Indonesia, kasus COVID-19 ini mulai berkembang sejak awal bulan Maret 2020 dan hingga detik ini masih terus terjadi peningkatan infeksi baru. Tujuan penelitian ini adalah melakukan prediksi pandemi COVID-19 dengan menggunakan model Random Forest dengan dataset COVID-19 di Indonesia mulai dari Maret sampai Desember 2020. Tahap pelaksanaan penelitian mengacu pada metode OSEMN yaitu Obtaining Data, Scrubbing Data, Exploring Data, Modeling Data, dan Interpreting data. Model Random Forest adalah salah satu algoritma machine learning yang sangat baik dalam memecahkan masalah klasifikasi dan regresi. Berdasarkan hasil penelitian didapatkan bahwa model yang dihasilkan memiliki tingkat akurasi 67 % di Indonesia dan nilai MAPE sebesar 33%. Sedangkan, untuk Sumatera Utara memiliki tingkat akurasi 98,4% dengan nilai MAPE sebesar 1,6%. Parameter terbaik dari model dengan nilai maxnodes 100 dan ntree 900 serta adanya trend kenaikan jumlah kasus COVID-19 beberapa waktu ke depan.

Kata kunci: Peramalan, Pandemi COVID-19, Random Forest, MAPE

Abstract

The COVID-19 pandemic that emerged in early 2020 has now spread throughout the world and affects all aspects of human life. In Indonesia, this COVID-19 case began to develop since the beginning of March 2020 and until now there is still an increase in new infections. The purpose of this research is to predict the COVID-19 pandemic using the Random Forest model with the COVID-19 dataset in Indonesia from March to December 2020. The research implementation stage refers to the OSEMN method, namely Obtaining Data, Scrubbing Data, Exploring Data, Data Modeling, and Data interpretation. The Random Forest Model is a machine learning algorithm that is excellent at solving classification and regression problems. Based on the research results, it was found that the resulting model has an accuracy rate of 67% in Indonesia and a MAPE value of 33%. Meanwhile, North Sumatra has an accuracy rate of 98.4% with a MAPE value of 1.6%. The best parameter of the model with a maxnodes value of 100 and ntree 900 as well as an increasing trend in the number of COVID-19 cases in the future.

Keywords: Forecasting, Pandemic COVID-19, Random Forest, MAPE

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Tuhan Yang Maha Esa atas segala rahmat-Nya sehingga penulis dapat melaksanakan dan menyelesaikan Tugas Akhir yang berjudul “Analisis Runtun Waktu Untuk Memprediksi Pandemi Covid-19 Dengan Model Random Forest” dengan baik dan tepat pada waktunya.

Adapun tujuan dari penulisan Tugas Akhir adalah untuk menyelesaikan pendidikan guna mendapatkan gelar Sarjana Strata Satu Program Studi Sistem Informasi pada Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Mikroskil Medan. Serta meningkatkan kemampuan berpikir kritis dan melakukan metode riset dengan lebih baik lagi.

Penulis menyadari Tugas Akhir ini tidak akan selesai tanpa adanya motivasi, bantuan, nasihat, bimbingan serta arahan dari berbagai pihak. Sehingga pada kesempatan ini penulis menyampaikan rasa terima kasih dan penghargaan yang sebesar-besarnya kepada:

1. Bapak Roni Yunis, S.Kom., M.T., selaku Dosen Pembimbing I yang bersedia meluangkan waktunya untuk memberikan bimbingan, pengarahan, dan semangat bagi penulis selama ini.
2. Ibu Angela, S.Kom., M.MSI, selaku Dosen Pendamping Pembimbing yang bersedia meluangkan waktunya untuk memberikan bimbingan dan pengarahan bagi penulis selama ini.
3. Bapak Dr. Pahala Sirait, S.T., M.Kom., selaku ketua STMIK Mikroskil Medan.
4. Ibu Rin Rin Meilani Salim, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi STMIK Mikroskil Medan.
5. Ibu Chatrine Sylvia, S.Kom., M.MSI, selaku Sekretaris Program Studi S-1 Sistem Informasi STMIK Mikroskil Medan.
6. Bapak Sudarto, S.Kom., M.Kom. selaku wali kelas Ebiz-B Pagi.
7. Bapak Hanes, S.Kom., M.Kom. selaku kepala Pusat Sistem Informasi.
8. Bapak-bapak dan Ibu-ibu Dosen STMIK Mikroskil Medan.
9. Teristimewa kedua Orang Tua Marchell Rianto yang telah memberikan dukungan dari berbagai sisi.

10. Teman-teman seperjuangan dalam penelitian yang sama-sama berjuang untuk menyelesaikan Tugas Akhir ini khususnya DS Research Team Bella Merlin Siregar, Ayu Ulfa Jamila, Fristiani Br. Sitepu, dan Vince Amelia Prada Sirait.
11. Teman diskusi yaitu alex yang selalu menyediakan waktu untuk berdiskusi dengan saya.

Penulis menyadari bahwa isi dan teknik penulisan Tugas Akhir ini masih jauh dari kesempurnaan baik dari tata bahasa maupun konsep materi yang terkandung di dalamnya yang disebabkan terbatasnya pengetahuan dan pengalaman penulis. Oleh karena itu kritik dan saran akan sangat diterima dengan senang hati untuk dapat dijadikan sebagai perbaikan untuk penulisan laporan selanjutnya.

Terima kasih

Medan, 1 Februari 2021

Penulis

Marchell Rianto

UNIVERSITAS MIKROSKIL

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup	3
1.4 Tujuan dan Manfaat.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 COVID-19	5
2.2 Peramalan (<i>Forecasting</i>)	9
2.2.1 Proses Peramalan (<i>Forecasting</i>)	10
2.2.2 Metode Peramalan.....	11
2.2.3 Langkah Sistem Peramalan	11
2.2.4 Jenis Peramalan.....	12
2.3 <i>Time Series</i>	13
2.3.1 Komponen <i>Time Series</i>	14
2.4 Analisis Runtun Waktu (<i>Time series analysis</i>)	14
2.4.1 Tipe pola data <i>Time Series</i>	15
2.4.2 Keuntungan Analisis Runtun Waktu.....	17
2.5 Stasioneritas.....	18
2.5.1 Konsep <i>Differencing</i>	19
2.5.2 Konsep <i>Augmented Dickey-Fuller (ADF)</i>	19
2.5.3 Pola grafik data Stasioner.....	21
2.6 Metode <i>Random Forest</i>	21
2.6.1 Perbedaan <i>Decision tree</i> dan <i>Random Forest</i>	23
2.6.2 <i>Random Forest</i> untuk peramalan <i>time series</i>	25

2.6.3	Implementasi <i>Random Forest</i>	26
2.7	Ketepatan Penggunaan Metode Peramalan	27
2.7.1	MAPE (<i>Mean Absolute Percent Error</i>)	28
2.8	Penelitian Terdahulu.....	29
BAB III	METODOLOGI PENELITIAN	33
3.1	Bahan Penelitian/Data	33
3.2	Peralatan Penelitian	34
3.3	Tahapan Penelitian dengan OSEM.....	34
3.3.1	<i>Obtaining Data</i> (Mengumpulkan data).....	35
3.3.2	<i>Scrubbing Data</i> (Pembersihan Data)	35
3.3.3	<i>Exploring Data</i> (Menjelajah Data).....	39
3.3.4	<i>Modelling Data</i> (Memodelkan Data).....	44
3.3.5	<i>Interpreting Data</i> (Menafsirkan Data).....	45
BAB IV	HASIL DAN PEMBAHASAN	47
4.1	Hasil Penelitian	47
4.1.1	Hasil Eksplorasi dan Analisis Data	47
4.1.2	Pemodelan prediksi dengan <i>Random Forest</i> untuk Indonesia	52
4.1.3	Pemodelan prediksi dengan <i>Random Forest</i> untuk Sumatera Utara.....	56
4.1.4	Hasil plot RMSE dan Variabel penting.....	59
4.2	Pembahasan	61
BAB V	KESIMPULAN DAN SARAN	62
5.1	Kesimpulan	62
5.2	Saran.....	62
DAFTAR PUSTAKA		63
LAMPIRAN		67
DAFTAR RIWAYAT HIDUP		71

DAFTAR GAMBAR

Gambar 2. 1 Analogi proses peramalan	10
Gambar 2. 2 Contoh grafik Time series	13
Gambar 2. 3 Model time series pada COVID-19.....	14
Gambar 2. 4 Contoh Plot data Trend	15
Gambar 2. 5 Contoh Plot data Seasonal.....	16
Gambar 2. 6 Contoh Plot data Horizontal.....	16
Gambar 2. 7 Contoh Plot data Cyclical.....	17
Gambar 2. 8 Plot runtun waktu stasioner.....	21
Gambar 2. 9 Plot runtun waktu non-stasioner.....	21
Gambar 2. 10 Analogi Random Forest seperti layaknya 2 pohon	22
Gambar 2. 11 Ilustrasi Decision tree.....	24
Gambar 2. 12 Ilustrasi Random Forest	24
Gambar 2. 13 Sampel test berubah dengan jumlah variabel prediktor	25
Gambar 2. 14 Implementasi Random Forest.....	27
Gambar 3. 1 Langkah-langkah OSEMN.....	35
Gambar 3. 2 Packages penting dalam scrubbing data.....	36
Gambar 3. 3 Pemanggilan Packages pada Scrubbing data	36
Gambar 3. 4 Memanggil dataset dan fungsi head()	37
Gambar 3. 5 Tampilan dataframe setelah beberapa variabel dihapus.....	37
Gambar 3. 6 Fungsi tail()	37
Gambar 3. 7 Mengidentifikasi kolom dengan fungsi names().....	38
Gambar 3. 8 Fungsi colnames()	38
Gambar 3. 9 Tampilan data sebelum variabel diubah.....	38
Gambar 3. 10 Tampilan data sesudah variabel diubah	38
Gambar 3. 11 Fungsi glimpse() membantu melihat tipe data	39
Gambar 3. 12 Mengubah tipe data date	39
Gambar 3. 13 Fungsi names() untuk melihat nama variabel	40
Gambar 3. 14 Fungsi glimpse().....	40
Gambar 3. 15 Fungsi head().....	41

Gambar 3. 16 Fungsi tail()	41
Gambar 3. 17 Fungsi dim, nrow dan ncol	41
Gambar 3. 18 Fungsi select()	41
Gambar 3. 19 Uji stasioner dengan adf.test().....	42
Gambar 3. 20 Fungsi summary()	43
Gambar 3. 21 Visualisasi dengan geom_line()	44
Gambar 4. 1 Visualisai data total kasus di Indonesia.....	47
Gambar 4. 2 Visualisasi data total meninggal di Indonesia	48
Gambar 4. 3 Visualisasi data total sembuh di Indonesia	49
Gambar 4. 4 Visualisasi data kasus baru di Indonesia	49
Gambar 4. 5 Visualisasi data kasus baru di Sumatera Utara.....	50
Gambar 4. 6 Visualisasi kasus di Indonesia.....	50
Gambar 4. 7 Visualisasi kasus di Sumatera Utara	51
Gambar 4. 8 Algoritma random forest dengan regresi pada Indonesia.....	52
Gambar 4. 9 Hasil pemrosesan prediksi di Indonesia	53
Gambar 4. 10 Visualisasi hasil prediksi total kasus meninggal di Indonesia	54
Gambar 4. 11 Visualisasi hasil prediksi total sembuh di Indonesia.....	55
Gambar 4. 12 Visualisasi hasil prediksi total kasus baru di Indonesia	55
Gambar 4. 13 Hasil MAPE prediksi Indonesia.....	56
Gambar 4. 14 Algoritma random forest dengan regresi pada SUMUT.....	56
Gambar 4. 15 Hasil pemrosesan prediksi di Sumatera Utara.....	57
Gambar 4. 16 Visualisasi hasil prediksi total kasus meninggal di Sumatera Utara... 58	
Gambar 4. 17 Visualisasi hasil prediksi total kasus sembuh di Sumatera Utara	58
Gambar 4. 18 Visualisasi hasil prediksi total kasus baru di Sumatera Utara.....	59
Gambar 4. 19 Hasil MAPE Sumatera Utara	59
Gambar 4. 20 Plot RMSE	60
Gambar 4. 21 Variabel Penting.....	61

DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu.....	29
Tabel 3. 1 Nama Variabel dan contoh data pada dataset.....	33



UNIVERSITAS MIKROSKIL