



# **MODEL PREDIKSI JUMLAH KASUS COVID-19 DAN ANALISIS BERBASIS DEMOGRAFI, SOSIAL DAN EKONOMI MENGGUNAKAN LONG SHORT TERM MEMORY (STUDI KASUS DI JAWA TIMUR)**

**AVRIENTA NOUVA EKA PUTRI**



**DEPARTEMEN ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2021**

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak mengurangi kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## **PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA**

Dengan ini saya menyatakan bahwa skripsi dengan judul “Model Prediksi Jumlah Kasus Covid-19 dan Analisis Berbasis Demografi, Sosial dan Ekonomi Menggunakan Long Short Term Memory (Studi Kasus di Jawa Timur)” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2021

*Avrienta Nouva Eka Putri*  
NIM G64170028

- Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak mengikuti kepentingan yang wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



AVRIENTA NOUVA EKA PUTRI. Model Prediksi Jumlah Kasus COVID-19 dan Analisis Berbasis Demografi, Sosial dan Ekonomi Menggunakan *Long Short Term Memory* (Studi Kasus di Jawa Timur). Dibimbing oleh YENI HERDIYENI dan MEDRIA KUSUMA DEWI HARDHIENATA.

*Coronavirus Disease 2019* atau (COVID-19) disebabkan oleh jenis virus corona baru yang dapat menimbulkan penyakit pernapasan. Penyakit ini ditemukan pertama kali di Wuhan, China, pada bulan Desember 2019. Virus ini ditularkan melalui cairan atau kontak langsung dan menginfeksi saluran pernapasan yang mengakibatkan pneumonia di sebagian besar kasus. Faktor demografi, sosial dan ekonomi merupakan faktor-faktor yang mendorong penyebaran penyakit menular. Antisipasi dan pengukuran dari berbagai faktor yang terlibat dalam penyakit menular dapat dimodelkan secara matematis dengan pemodelan matematika. Penelitian ini melakukan pemodelan prediksi kasus COVID-19 di Jawa Timur menggunakan *Long Short Term Memory* (LSTM). Berdasarkan hasil penelitian, kabupaten/kota di Jawa Timur terbagi menjadi lima *cluster* sesuai karakteristik demografi, sosial dan ekonominya. Jumlah kasus positif COVID-19 yang tinggi terdapat di *cluster* dengan PDRB yang tinggi. Hasil analisis regresi menunjukkan bahwa atribut PDRB merupakan atribut yang berpengaruh signifikan terhadap jumlah terkonfirmasi COVID-19 di Jawa Timur. Penelitian ini memperkirakan jumlah terkonfirmasi positif COVID-19 masih terus naik hingga akhir Juli 2021.

Kata Kunci: *COVID-19, demografi, ekonomi, long short term memory, pemodelan prediksi, sosial.*

- Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak mengujikan kepentingan wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



AVRIENTA NOUVA EKA PUTRI. *Prediction Model Number of COVID-19 Cases and Analysis Based on Demography, Social and Economy Using Long Short Term Memory (Case Study: East Java Province)*. Supervised by YENI HERDIYENI and MEDRIA KUSUMA DEWI HARDHIENATA.

*Coronavirus Disease 2019 or (COVID-19) is caused by a new type of coronavirus that can cause respiratory disease. This disease was first discovered in Wuhan, China, in December 2019. This virus is transmitted through fluids or direct contact and infects the respiratory tract resulting in pneumonia in most cases . Demographic, social and economic factors are factors that encourage the spread of infectious diseases. The anticipation and measurement of various factors involved in infectious diseases can be modeled mathematically. This research conducts predictive modeling of COVID-19 cases in East Java using Long Short Term Memory (LSTM). Based on the research results, districts/cities in East Java are divided into five clusters according to their demographic, social and economic characteristics. Many positive cases of COVID-19 are found in clusters with a high value of Gross Regional Domestic Product (GDP regional). The results of the regression analysis indicate that GDP regional attribute is an attribute that has a significant effect on the number of positive confirmed cases of COVID-19 in East Java. This study estimates that the addition of the positive number of COVID-19 will continue to rise until the end of July 2021.*

**Keywords:** COVID-19, demographic, economic, long short term memory, prediction modeling, social.

- Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

# IPB University

©Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.





## ©Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah,
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2021

Hak Cipta dilindungi Undang-Undang

*Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.*

*Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.*



## **MODEL PREDIKSI JUMLAH KASUS COVID-19 DAN ANALISIS BERBASIS DEMOGRAFI, SOSIAL DAN EKONOMI MENGGUNAKAN LONG SHORT TERM MEMORY (STUDI KASUS DI JAWA TIMUR)**

**AVRIENTA NOUVA EKA PUTRI**

Skripsi  
sebagai salah satu syarat untuk memperoleh gelar  
Sarjana pada  
Program Studi Ilmu Komputer

**DEPARTEMEN ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2021**



## ©Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## IPB University

Tim Penguji pada Ujian Skripsi:  
Dr.Eng. Wisnu Ananta Kusuma, S.T., M.T.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Model Prediksi Kasus Covid-19 dan Analisis Berbasis Demografi, Sosial dan Ekonomi Menggunakan Long Short Term Memory (Studi Kasus di Jawa Timur)

Nama : Avrienta Nouva Eka Putri  
NIM : G64170028

Disetujui oleh



Pembimbing 1:  
Dr. Yeni Herdiyeni, S.Si., M.Kom.

Pembimbing 2:  
Dr. Medria K. D. Hardhienata, S.Komp.

Diketahui oleh



Ketua Departemen Ilmu Komputer:  
Dr. Sony Hartono Wijaya, S.Kom., M.Kom.  
NIP 19810809 200812 1 002

Tanggal Ujian:  
2 Agustus 2021

Tanggal Lulus:



- Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak mengurangi kepentingan wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhanaahu wa ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan dengan judul Model Prediksi Jumlah Kasus Covid-19 dan Analisis Berbasis Demografi, Sosial dan Ekonomi Menggunakan *Long Short Term Memory* (Studi Kasus di Jawa Timur).

Terima kasih penulis ucapan kepada dosen pembimbing Dr. Yeni Herdiyeni, S.Si., M.Kom dan Dr. Medria K. D. Hardhienata, S.Komp yang telah membimbing dan banyak memberi saran. Terimakasih juga penulis ucapan kepada dosen pengaji Dr.Eng. Wisnu Ananta Kusuma, S.T., M.T yang telah memberikan nasihat dan saran agar penelitian saya menjadi lebih baik. Ungkapan terima kasih juga disampaikan penulis kepada orang tua tercinta yaitu Bapak Avries Praputro Wulandoro, Ibu Swarini Wigati, serta seluruh keluarga yang telah memberikan dukungan, doa, dan kasih sayangnya. Teman-teman Ilmu Komputer 54 yang sudah bersama selama 4 tahun ini. *Last but not least, I wanna thank me, for believing in me, for doing all this hard work, for having no days off, for never quitting, for just being me at all times.*

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2021

*Avrienta Nouva Eka Putri*



<b>DAFTAR ISI</b>	11
<b>DAFTAR TABEL</b>	13
<b>DAFTAR GAMBAR</b>	14
<b>I PENDAHULUAN</b>	14
1.1 Latar Belakang	14
1.2 Perumusan Masalah	16
1.3 Tujuan	16
1.4 Manfaat	16
1.5 Ruang Lingkup	16
1.6 Hipotesis	17
<b>II TINJAUAN PUSTAKA</b>	17
2.1 Kondisi Demografi, Sosial dan Ekonomi di Jawa Timur	17
2.2 Coronavirus Disease (COVID-19)	17
2.3 Long Short Term Memory	17
2.4 K-Means <i>Clustering</i>	18
<b>III METODE</b>	21
3.1 Tahapan Penelitian	21
3.2 Data Penelitian	21
3.3 Pengumpulan Data	22
3.4 Praproses dan Eksplorasi Data	22
3.5 Analisis Regresi	23
3.6 Analisis <i>Clustering</i>	24
3.7 Pemodelan Prediksi Menggunakan LSTM	24
3.8 Evaluasi Model	24
3.9 Lingkungan Pengembangan	24
<b>IV HASIL DAN PEMBAHASAN</b>	25
4.1 Praproses dan Eksplorasi Data	25
4.2 Analisis Regresi	27
4.3 Analisis <i>Clustering</i>	28
4.4 Pemodelan Prediksi Menggunakan LSTM	32
4.5 Evaluasi Model	33
<b>V SIMPULAN DAN SARAN</b>	35
5.1 Kesimpulan	35
5.2 Saran	35
<b>DAFTAR PUSTAKA</b>	36

Hak Cipta Dilindungi Undang-undang  
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
 b. Pengutipan tidak mengugikan kepentingan yang wajar IPB University.  
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## RIWAYAT HIDUP

39

- Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

12



<b>DAFTAR TABEL</b>	
Tabel 1 Atribut data kasus COVID-19 dan data demografi, sosial dan ekonomi di tiap kabupaten/kota di Jawa Timur	21
Tabel 2 Kategorisasi data kasus konfirmasi dan meninggal COVID19	26
Tabel 3 Normalisasi data demografi, sosial, ekonomi	26
Tabel 4 Nilai VIF setelah menghilangkan beberapa atribut	27
Tabel 5 Koefisien Regresi Masing-Masing Variabel Terpilih	28
Tabel 6 <i>Cluster</i> , Wilayah, Status Covid dan Status Fatality Rate	29
Tabel 7 Nilai rata-rata atribut tiap <i>cluster</i>	30
Tabel 8 Struktur LSTM	33

**DAFTAR GAMBAR**

Gambar 1 Contoh arsitektur LSTM (Marinescu 2019)	19
Gambar 2 Tahapan Penelitian	21
Gambar 3 Jumlah konfirmasi COVID19 di Jawa Timur pada bulan Maret 2020 sampai Juni 2021	25
Gambar 4 Wilayah kabupaten/kota di Jawa Timur dengan jumlah kasus konfirmasi COVID19 terendah hingga tertinggi	26
Gambar 5 Korelasi Pearson's data demografi, sosial dan ekonomi	27
Gambar 6 Jumlah <i>Cluster</i> yang Optimal	29
Gambar 7 Grafik perbandingan rataan per atribut setiap <i>cluster</i>	31
Gambar 8 Plot perbandingan nilai aktual dan prediksi kasus konfirmasi COVID19	33

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak mengujikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## 1.1 Latar Belakang

*Coronavirus Disease 2019* atau (COVID-19) disebabkan oleh jenis virus corona baru yang dapat menimbulkan penyakit pernapasan. Penyakit ini diberitahukan pertama kali di Wuhan, China pada Desember 2019 (Cheng *et al.* 2020). Virus COVID-19 ditularkan melalui cairan atau kontak langsung dan menginfeksi saluran pernapasan yang mengakibatkan pneumonia di sebagian besar kasus (Ragab *et al.* 2020). Per tanggal 1 November 2020, *World Health Organization* (2020) telah melaporkan secara global 45.968.799 jumlah kasus yang dikonfirmasi dan 1.192.911 jumlah kematian. Pada tanggal 15 November 2020, berdasarkan data Worldometers (2020), Indonesia menempati urutan ke 21 dengan jumlah kasus COVID-19 tertinggi di dunia dengan jumlah kasus yang dikonfirmasi mencapai 467.113 dan jumlah kasus kematian 15.211. Salah satu sumber utama penularan virus COVID-19 adalah di kota-kota metropolitan, salah satunya adalah kota yang berada di Provinsi Jawa Timur yaitu Surabaya, yang merupakan kota metropolitan terbesar kedua di Indonesia (Turista *et al.* 2020). Provinsi Jawa Timur sendiri menempati urutan kedua setelah Provinsi DKI Jakarta dengan jumlah kasus COVID-19 tertinggi di Indonesia yaitu 56.070 kasus (KPCPEN 2020).

Penting untuk memahami faktor yang terlibat dalam keberhasilan penyebaran Sars-Cov-2 ini dan memerlukan identifikasi dan kuantifikasi dari faktor perilaku, sosial-ekonomi dan lingkungan yang memungkinkan penularan infeksi antara manusia (Stein *et al.* 2014). Faktor utama yang berpengaruh dalam penyebaran patogen adalah kepadatan populasi manusia yang berperan sebagai penggerak sosial-ekonomi, kemudian diikuti oleh faktor lainnya yaitu kondisi ekologi atau lingkungan (Guernier *et al.* 2004). Faktor ekologi, lingkungan, atau demografis juga merupakan faktor yang mendorong penyebaran penyakit menular. Faktor-faktor ini semakin meningkat prevalensinya. Peningkatan ini menunjukkan bahwa penyakit infeksi akan terus dan mungkin meningkat serta menekankan kebutuhan mendesak untuk pengawasan dan pengendalian yang efektif (Morse 1995).

Antisipasi dan pengukuran dari berbagai faktor yang terlibat dalam penyakit menular dapat dimodelkan secara matematis dengan pemodelan matematika (Heesterbeek *et al.* 2015). Model matematika dapat digunakan untuk menghubungkan proses biologis penularan dan dinamika infeksi yang muncul di tingkat populasi (Grassly 2008). Pengembangan model epidemiologi memiliki kompleksitas yang cukup tinggi, oleh karena itu, pembelajaran *machine learning* baru-baru ini mendapat perhatian untuk membangun model prediksi penyakit menular. Pendekatan *machine learning* bertujuan mengembangkan model dengan kemampuan generalisasi yang lebih tinggi dan keandalan prediksi yang lebih besar (Ardabili *et al.* 2020).

Metode *machine learning* yang akan digunakan untuk membangun model prediksi COVID19 pada penelitian ini adalah *Long Short Term Memory* (LSTM). Karena fluktuatifnya data penambahan kasus COVID19, metode LSTM dinilai cukup berhasil untuk meminimalkan error (Adhinata dan Rakhmadani 2021). Selain itu metode LSTM merupakan model terbaik untuk data dinamis yaitu data

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak mengugik kepentingan wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



yang tiba-tiba berubah drastis dalam berbagai kasus (Bodapati *et al.* 2020).

Penelitian Liu *et al.* (2018) membuat prediksi tren influenza menggunakan LSTM menggunakan data pengawasan virologi, penyebaran geografis influenza, *Google trends*, iklim dan polusi udara. Korelasi pearson juga digunakan untuk membuktikan PM2.5 dan karbon monoksida (CO) berpengaruh signifikan terhadap laju *influenza like illness* (ILI). Terbukti bahwa PM2.5 dan CO berkorelasi signifikan dengan laju ILI. Peneliti juga menguji 4 model LSTM yang berbeda yaitu ukuran sampel besar/kecil yang variabel predikturnya berasal dari polusi udara, perubahan suhu, dan set data *Google trends* dan ukuran sampel besar/kecil yang hanya menggunakan data *Google trends* sebagai variabel prediktor. Hasil penelitian ini menunjukkan bahwa metode penggunaan sumber data multivariat untuk tren influenza lebih baik daripada metode yang hanya menggunakan data *Google trends*. Selain itu, penelitian ini menemukan bahwa ukuran sampel yang besar dapat meningkatkan model prediksi influenza dengan jelas. Penelitian Chae *et al.* (2020) telah membuat model prediksi 3 penyakit menular (*chickenpox*, *scarlet fever* dan *malaria*) satu minggu kedepan menggunakan model pembelajaran *deep neural network* (DNN) dan *long short term memory* (LSTM) dan dibandingkan dengan *autoregressive integrated moving average* (ARIMA). Hasil penelitian menunjukkan bahwa model DNN dan LSTM memiliki kinerja yang lebih baik daripada ARIMA. Saat memprediksi *chickenpox*, 10 model DNN dan LSTM teratas meningkatkan kinerja rata-rata masing-masing sebesar 24% dan 19%. Model DNN bekerja secara stabil dan model LSTM dinilai lebih akurat untuk memprediksi penyebaran penyakit menular.

## 2.2 Perumusan Masalah

Berdasarkan latar belakang tersebut, maka penulis merumuskan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana menganalisis variabel Demografi, Sosial dan Ekonomi wilayah Jawa Timur dan melihat hubungan variabel tersebut dengan jumlah kasus COVID-19
2. Bagaimana cara menerapkan algoritma *long short term memory* (LSTM) untuk memprediksi kasus COVID-19 di Jawa Timur?

## 2.3 Tujuan

Tujuan dari penelitian ini adalah:

1. Menganalisis dan melihat hubungan variabel Demografi, Sosial dan Ekonomi terhadap jumlah kasus COVID-19 di tiap kabupaten/kota Jawa Timur
2. Memprediksi jumlah kasus COVID-19 menggunakan metode *Long Short Term Memory* (LSTM).

## 2.4 Manfaat

Hasil penelitian ini diharapkan dapat membantu pihak pemerintah Provinsi Jawa Timur untuk menentukan kebijakan penanganan penyebaran COVID-19 di Jawa Timur



## 2.5 Ruang Lingkup

Ruang lingkup penelitian ini adalah:

1. Data yang digunakan adalah data kejadian COVID-19 dari tanggal 20 Maret 2020 sampai 30 Juni 2021
2. Data Demografi, Sosial, Ekonomi pada tahun 2020.
3. Daerah yang menjadi fokus pada penelitian ini adalah kabupaten/kota Provinsi Jawa Timur.

## 2.6 Hipotesis

Adanya pengaruh yang cukup signifikan dari variabel PDRB serta daya beli atau konsumsi masyarakat pada tiap *cluster* di wilayah kabupaten/kota Jawa Timur terhadap penambahan kasus terkonfirmasi COVID-19

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.



## II TINJAUAN PUSTAKA

### 2.1 Kondisi Demografi, Sosial dan Ekonomi di Jawa Timur

Berdasarkan BPS (2019) jumlah penduduk Jawa Timur hingga tahun 2019 mencapai 39.698.000 jiwa sehingga menjadikan Jawa Timur sebagai Provinsi dengan populasi tertinggi di Indonesia kedua setelah Jawa Barat. Faktor demografi seperti angka kelahiran dan kematian mempengaruhi pertumbuhan populasi penduduk (Bavel 2013). Jawa Timur sendiri merupakan daerah kedua setelah DKI Jakarta dengan kontribusi terbesar di wilayah Jawa pada perekonomian nasional yaitu sebesar 14.63%. Pertumbuhan ekonomi diperkirakan akan mengalami penurunan akibat dari penurunan pendapatan masyarakat dan aktivitas ekonomi yang juga terjadi karena pembatasan sosial berskala besar (PSBB). Terlebih lagi, PSBB diterapkan di banyak wilayah yang mempunyai kontribusi ekonomi besar, salah satunya adalah Jawa Timur (BPS 2020). Salah satu indikator utama sosial untuk mengukur keberhasilan dalam upaya membangun kualitas hidup adalah Indeks Pembangunan Manusia (IPM). Pada tahun 2018, IPM Jawa Timur tumbuh 0.72% dengan rata-rata pertumbuhan IPM mencapai 0.98% per tahun (Pemprov Jawa Timur 2019).

### 2.2 Coronavirus Disease (COVID-19)

*Coronavirus Disease 2019* atau (COVID-19) disebabkan oleh jenis virus corona baru yang menyebabkan wabah penyakit pernapasan diberitahukan pertama kali di Wuhan, China pada Desember 2019 (Cheng *et al.* 2020). Penyakit pernapasan yang baru muncul ini disebabkan oleh sindrom pernapasan akut parah *coronavirus* 2 (SARS-CoV-2), baru-baru ini menjadi pandemi (Cao 2020). Sebagian besar pasien dengan COVID-19 menunjukkan gejala ringan hingga sedang, tetapi sekitar 15% berkembang menjadi pneumonia berat dan sekitar 5% berkembang menjadi acute respiratory distress syndrome (ARDS), syok septik atau kegagalan beberapa organ (Cao 2020). Gejala umum yang biasanya muncul saat terinfeksi virus COVID-19 adalah demam, batuk, kelelahan, sesak nafas dan sakit kepala (Tian *et al.* 2020). Berdasarkan WHO (2020) vaksin COVID-19 belum juga tersedia. Banyak vaksin potensial untuk COVID-19 yang masih sedang dipelajari dan dilakukan beberapa uji klinis besar.

### 2.3 Long Short Term Memory

Jaringan *Long Short-Term Memory* (LSTM) adalah recurrent neural networks yang dilengkapi dengan mekanisme *gating* khusus yang mengontrol akses ke sel memori (Hochreiter dan Schmidhuber 1997). LSTM mampu menangani penghafalan dan pengingatan kembali untuk jangka panjang, khususnya data yang sangat besar. LSTM pada prinsipnya dapat menggunakan unit memorinya untuk mengingat informasi yang jaraknya jauh dan melacak berbagai atribut teks yang sedang diproses (Karpathy *et al.* 2016). Secara khusus, LSTM menggunakan 3 vektor *gate* yaitu *input gate*, *forget gate*, dan *output gate*. Vektor *input gate* mengontrol seberapa banyak vektor masukan yang diizinkan untuk mempengaruhi vektor memori. Vektor *forget gate* mengontrol seberapa

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak mengurangi kepentingan yang wajar IPB University.

banyak vektor memori sebelumnya yang harus dilupakan. Vektor *output gate* mengontrol berapa banyak vektor memori yang disimpan dalam *hidden state* (Kalchbrenner *et al.* 2016).

Persamaan variabel dari LSTM yaitu  $x_t$  sebagai input vektor ke LSTM,  $F_t$  sebagai vektor aktivasi *forget gate*,  $I_t$  sebagai vektor aktivasi *input gate*,  $O_t$  sebagai vektor aktivasi *output gate*,  $H_t$  sebagai vektor *hidden state* atau dikenal sebagai vektor keluaran dari unit LSTM,  $C_t$  sebagai vektor *cell state*,  $W-U-b$  sebagai matriks bobot dan parameter vektor bias yang perlu dipelajari selama pelatihan. Bentuk persamaan LSTM (Zhang *et al.* 2016) dapat dilihat pada persamaan (1), (2), (3), (4) dan (5).

$$F_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (1)$$

$$I_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

$$O_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$C_t = F_t * c_{t-1} + I_t * \sigma_g(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

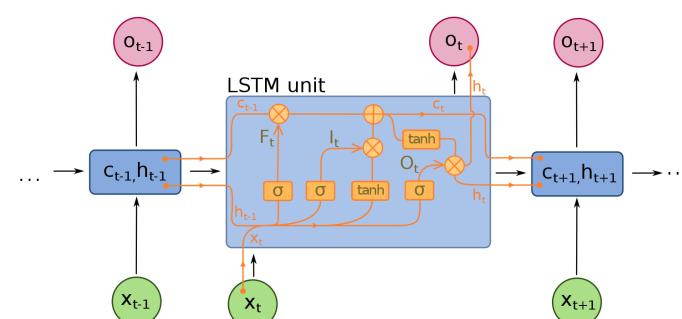
$$H_t = O_t * \sigma_h(c_t) \quad (5)$$

LSTM menggunakan dua fungsi aktivasi, yaitu fungsi aktivasi sigmoid dan tanh. Fungsi aktivasi sigmoid akan mengubah suatu nilai  $x$  menjadi nilai yang memiliki rentang 0 sampai 1. Fungsi aktivasi tanh akan mengubah suatu nilai  $x$  menjadi nilai yang memiliki rentang -1 sampai 1. Rumus fungsi aktivasi sigmoid dan tanh (Karlic dan Olgac 2016) dapat dilihat pada persamaan (6) dan (7).

$$\sigma = \frac{1}{1+e^{-x}} \quad (6)$$

$$\tanh(x) = 2\sigma(2x) - 1 \quad (7)$$

dengan nilai  $x$  adalah data input dan  $\sigma$  adalah nilai dari fungsi aktivasi sigmoid. Contoh arsitektur LSTM dapat dilihat pada Gambar 1 (Marinescu 2019).



Gambar 1 Contoh arsitektur LSTM (Marinescu 2019)

## 2.4 K-Means Clustering

*K-Means* adalah metode *clustering* berbasis jarak yang membagi data ke dalam sejumlah *cluster* (Mahmudan 2020). *K-Means* sangat terkenal karena kemudahan dan kemampuannya mengelompokkan data yang besar dan data *outlier* dengan sangat cepat. Data-data dipilih menjadi beberapa kelompok dengan



kriteria yang telah ditentukan lalu dikumpulkan menjadi satu dalam sebuah *cluster*. Dimana setiap *cluster* memiliki titik pusat yang disebut *centroid*. Berikut adalah tahapan-tahapan untuk melakukan *K-Means clustering*:

1. Menentukan banyak  $k$  *cluster* yang ingin dibentuk. Pusat *cluster* awal (*centroid*) ditentukan sebarang objek secara random
2. Menghitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus *Eucledian Distance* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Rumus *Eucledian Distance* dapat dilihat pada persamaan berikut

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \quad (8)$$

dengan  $d(x_i, \mu_j)$  adalah jarak antara *cluster*  $x$  dengan pusat *cluster*  $\mu$  pada kata ke  $i$ ,  $x_i$  adalah bobot ke  $i$  pada *cluster* yang ingin dicari jaraknya,  $\mu_i$  bobot kata ke  $i$  pada pusat *cluster*

3. Menghitung nilai *centroid* baru. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan persamaan berikut

$$C_k = \frac{1}{n_k} \sum di \quad (9)$$

dengan  $n_k$  adalah jumlah data dalam *cluster* dan  $di$  adalah jumlah dari nilai jarak yang masuk dalam masing-masing *cluster*

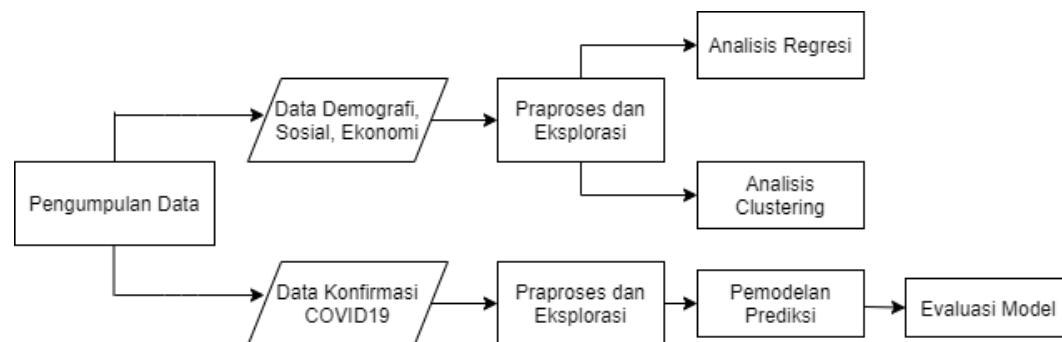
4. Menghitung jarak setiap objek dengan tiap *centroid* yang baru, sama seperti langkah 2
5. Mengulangi langkah 2 dan 3 sampai tidak ada perubahan anggota cluster. Jika telah terpenuhi, maka nilai rata-rata pusat *cluster* ( $\mu_j$ ) pada iterasi

terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data

## III METODE

### 3.1 Tahapan Penelitian

Tahapan penelitian dapat dilihat pada Gambar 2. Tahapan ini terdiri dari beberapa tahapan yaitu pengumpulan data, praproses dan eksplorasi data, analisis regresi, analisis *clustering*, pemodelan prediksi dan evaluasi model.



Gambar 2 Tahapan Penelitian

### 3.2 Data Penelitian

Data yang akan digunakan pada penelitian ini terbagi menjadi dua, yaitu data kasus konfirmasi dan meninggal COVID-19 di kota/kabupaten Provinsi Jawa Timur dan data demografi, sosial dan ekonomi di kota/kabupaten Provinsi Jawa Timur. Atribut yang akan digunakan pada data kasus COVID-19 sebanyak 2 atribut sedangkan pada data demografi, sosial dan ekonomi sebanyak 12 atribut. Total atribut yang akan digunakan sebanyak 14 atribut.

Tabel 1 Atribut data kasus COVID-19 dan data demografi, sosial dan ekonomi di tiap kabupaten/kota di Jawa Timur

Nama Atribut	Jenis Data	Keterangan
Konfirmasi	Time Series, Numerik	Jumlah kasus terkonfirmasi COVID-19
Meninggal	Time Series, Numerik	Jumlah kasus meninggal karena COVID-19
Jumlah Rumah Sakit Umum	Numerik	Banyaknya rumah sakit umum yang ada di suatu daerah
Jumlah Dokter	Numerik	Banyaknya dokter yang ada di suatu daerah
Indeks Pembangunan Manusia (IPM)	Numerik	Indikator untuk mengukur keberhasilan dalam upaya membangun kualitas hidup masyarakat
Persentase Penduduk Miskin	Numerik	Persentase penduduk yang berada dibawah Garis Kemiskinan



Produk Domestik Numerik Regional Bruto (PDRB)	Numerik	Jumlah nilai tambah atas barang dan jasa yang dihasilkan oleh berbagai unit produksi di tiap kabupaten/kota di Jawa Timur dalam jangka waktu tertentu (biasanya satu tahun).
Total Pengeluaran Numerik Perkapita	Numerik	Biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga yang telah disesuaikan dengan paritas daya beli
Tingkat Pengangguran Terbuka	Numerik	Persentase jumlah pengangguran terhadap jumlah angkatan kerja
Kepadatan Penduduk	Numerik	Rasio banyaknya penduduk per kilometer persegi
Jumlah Populasi	Numerik	Total orang yang sudah menetap di suatu wilayah paling sedikit 6 bulan atau kurang dari 6 bulan tetapi bermaksud untuk menetap
Persentase Penduduk Perkotaan	Numerik	Perbandingan jumlah penduduk perkotaan terhadap total seluruh penduduk yang ada pada daerah tersebut
Rasio Jenis Kelamin	Numerik	Perbandingan antara jumlah penduduk pria dan jumlah penduduk wanita pada suatu daerah dan pada waktu tertentu
Persentase Penduduk Lansia	Numerik	Perbandingan jumlah penduduk usia 60 tahun ke atas dengan jumlah seluruh penduduk

### 3.3 Pengumpulan Data

Data yang akan digunakan pada penelitian ini terbagi menjadi dua, yaitu data kasus konfirmasi COVID-19 di kota/kabupaten Provinsi Jawa Timur yang akan digunakan untuk membuat model prediksi, dan data demografi, sosial dan ekonomi yang akan digunakan untuk analisis regresi dan analisis *clustering*. Data kasus konfirmasi COVID-19 diambil dari data yang telah dikumpulkan oleh pemerintah Provinsi Jawa Timur melalui halaman web [infocovid19.jatimprov.go.id](http://infocovid19.jatimprov.go.id), sedangkan untuk data demografi, sosial dan ekonomi diambil dari Badan Pusat Statistik Provinsi Jawa Timur pada tahun 2020.

### 3.4 Praproses dan Eksplorasi Data

Pada tahap ini dilakukan pengecekan *missing value*, kategorisasi data berdasarkan mean dan standar deviasi, menghilangkan *multicollinearity* dan normalisasi data. Pada tahap kategorisasi data, tingkat keparahan kasus COVID-19 di setiap kabupaten/kota dibagi menjadi lima kategori yaitu sangat rendah, rendah, sedang, tinggi dan sangat tinggi berdasarkan kasus kumulatif positif COVID-19 pada tanggal 30 Juni 2021. Yang menjadi nilai batas pada kategorisasi adalah nilai rata-rata ( $\bar{x}$ ) dan simpangan baku ( $S$ ) kasus positif COVID-19. Proses kategorisasi dilakukan sebagai berikut:

- Sangat tinggi. Kasus COVID-19 dikategorikan sangat tinggi jika:



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- Pengutipan tidak merugikan kepentingan wajar IPB University.

- $x \geq \bar{x} + (1,5 \times S)$
- b. Tinggi. Kasus COVID-19 dikategorikan tinggi jika:  
 $\bar{x} + (0,5 \times S) \leq x < \bar{x} + (1,5 \times S)$
- c. Sedang. Kasus COVID-19 dikategorikan sedang jika:  
 $\bar{x} - (0,5 \times S) \leq x < \bar{x} + (0,5 \times S)$
- d. Rendah. Kasus COVID-19 dikategorikan rendah jika:  
 $\bar{x} - (1,5 \times S) \leq x < \bar{x} - (0,5 \times S)$
- e. Sangat rendah. Kasus COVID-19 dikategorikan sangat rendah jika:  
 $x < \bar{x} - (1,5 \times S)$

Pada tahap menghilangkan multicollinearity, dilakukan dengan beberapa cara untuk memeriksa multikolinearitas di antaranya adalah dengan menghitung nilai korelasi Pearson antar variabel atau menghitung nilai Variance Inflation Factor (VIF). Koefisien korelasi Pearson mengukur kekuatan hubungan linear antara dua variabel. Nilai yang mendekati 1 atau -1 menunjukkan bahwa terdapat hubungan yang kuat antara kedua variabel tersebut. Setelah menghitung nilai koefisien korelasi Pearson, berikutnya dilakukan perhitungan nilai Variance Inflation Factor (VIF) pada setiap atribut. Nilai  $VIF < 10$  menunjukkan bahwa terdapat korelasi yang kuat antar variabel bebas. Salah satu cara yang dapat digunakan untuk menghilangkan multikolinearitas yaitu dengan tidak mengikutsertakan atribut yang saling berkorelasi kuat tersebut pada proses analisis berikutnya.

Pada tahap normalisasi data, data yang memiliki nilai besar diskalakan ke dalam rentang nilai kecil. Metode normalisasi yang akan digunakan dalam penelitian ini adalah min-max dengan rentang 0-1. Rumus normalisasi min-max (Liping *et al.* 2009) dapat dilihat pada persamaan (10):

$$y_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (10)$$

Keterangan :

- $y_{ij}$  = normalisasi  
 $x_{ij}$  = nilai yang dinormalisasi  
 $\min(x_j)$  = nilai minimum data  
 $\max(x_j)$  = nilai maksimum data

### 3.5 Analisis Regresi

Analisis Regresi dilakukan untuk melihat hubungan sebab-akibat antara variabel independen dengan variabel dependen (Aiken *et al.* 2012). Analisis regresi dilakukan untuk mendapatkan koefisien regresi yang menunjukkan pengaruh variabel independen yaitu variabel demografi, sosial dan ekonomi terhadap variabel dependen yaitu jumlah kasus terkonfirmasi COVID19. Analisis regresi ini menggunakan algoritma *Multiple Linear Regression*. Semakin tinggi nilai koefisien regresi, maka semakin tinggi juga pengaruh variabel independen terhadap variabel dependen.



### 3.6 Analisis *Clustering*

Pada tahap ini data demografi, sosial, ekonomi dilakukan *clustering* menggunakan K-Means dengan tujuan untuk melihat karakteristik wilayah berdasarkan faktor demografi, sosial dan ekonomi. *K-Means clustering* merupakan salah satu algoritma pengelompokan yang digunakan dalam kelompok *unsupervised learning* yang digunakan untuk mengelompokkan data menjadi beberapa kelas dengan partisi sistem (Hutagalung *et al.* 2020)

### 3.7 Pemodelan Prediksi Menggunakan LSTM

Pemodelan dilakukan dengan menggunakan LSTM. Keluaran dari proses *learning* ini adalah hasil prediksi jumlah kasus konfirmasi COVID-19. LSTM akan menggunakan persamaan (1), (2), (3), (4), dan (5) untuk menghitung vektor pada setiap gates. Nilai  $x_t$  merupakan sebuah *input* data berupa data kasus konfirmasi COVID-19 pada hari sebelumnya yang nilainya sudah diubah menjadi sebuah vektor. Selain itu diperlukan fungsi aktivasi untuk membangkitkan *output* dan inisialisasi beberapa parameter

### 3.8 Evaluasi Model

Hasil model prediksi dievaluasi dengan melihat nilai *root mean squared error* (RMSE). RMSE merupakan metode evaluasi untuk mengukur perbedaan antara nilai aktual dan nilai prediksi, metode evaluasi ini banyak digunakan dalam mengevaluasi prediksi penyakit menular (Chae *et al.* 2020). Nilai RMSE dapat dihitung menggunakan persamaan (11)

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (X_t - \hat{X}_t)^2}{n}} \quad (11)$$

Keterangan :

$n$  = jumlah data yang diprediksi

$X$  = nilai jumlah kasus COVID-19 aktual

$\hat{X}$  = nilai hasil prediksi COVID-19

### 3.9 Lingkungan Pengembangan

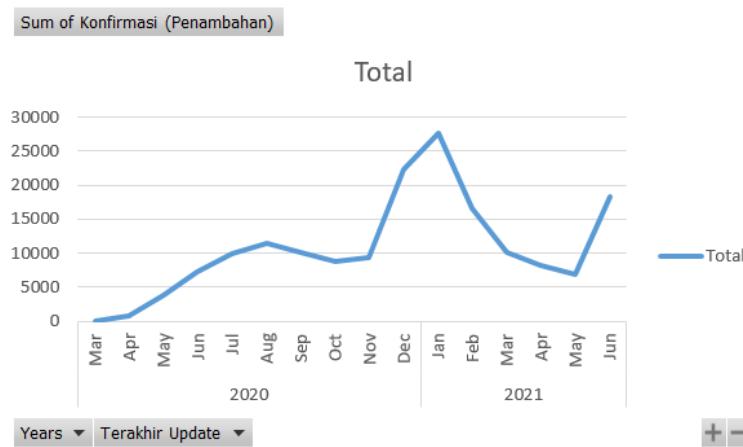
Spesifikasi perangkat keras dan perangkat lunak yang digunakan untuk penelitian ini adalah sebagai berikut:

1. Perangkat keras berupa laptop dengan spesifikasi:
  - Prosesor Intel Core i5-7200U CPU 2.50 GHz
  - RAM 8 GB
2. Perangkat lunak yang digunakan, yaitu:
  - Sistem operasi Windows 10 64-bit
  - Bahasa Pemrograman Python 3.7
  - Google Collab untuk pemanfaatan GPU pada tahapan pemodelan

## IV HASIL DAN PEMBAHASAN

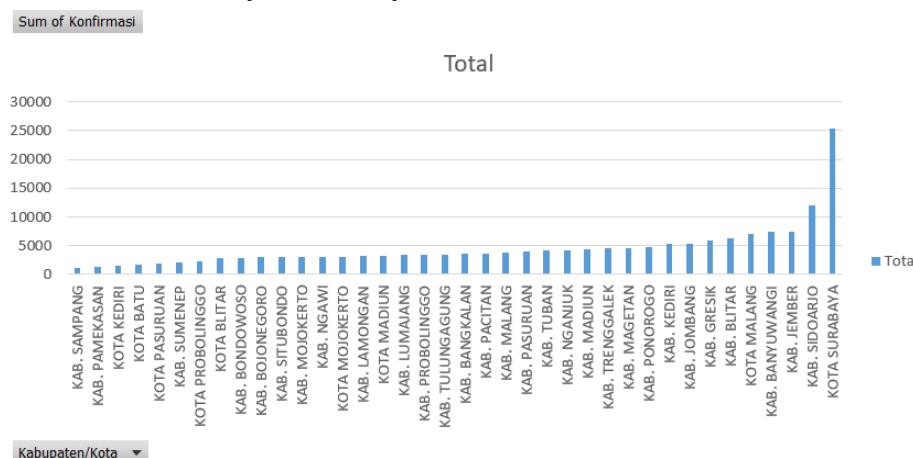
#### **4.1 Praproses dan Eksplorasi Data**

Pada tahap eksplorasi data terlihat bahwa, terjadi peningkatan penambahan kasus terkonfirmasi COVID-19 di provinsi Jawa Timur yang cukup tajam di bulan Desember 2020 dan Januari 2021 dan di Bulan Februari 2021 dan Mei 2021 jumlah kasus terkonfirmasi mulai terjadi penurunan kemudian terjadi kenaikan yang cukup drastis pada Bulan Juni 2021.



Gambar 3 Jumlah konfirmasi COVID19 di Jawa Timur pada bulan Maret 2020 sampai Juni 2021

Per tanggal 30 Juni 2021 Kota Surabaya merupakan kota dengan jumlah kasus terkonfirmasi tertinggi di Jawa Timur yaitu sebanyak 25397 kasus dan Kabupaten Sampang merupakan kabupaten dengan jumlah kasus terkonfirmasi terendah di Jawa Timur yaitu sebanyak 1221 kasus





Gambar 4 Wilayah kabupaten/kota di Jawa Timur dengan jumlah kasus konfirmasi COVID19 terendah hingga tertinggi

Pada tahap praproses, data jumlah konfirmasi dan meninggal pada tanggal 30 Juni 2021 di setiap wilayah kabupaten/kota Jawa Timur dikategorisasikan menjadi sangat tinggi, tinggi, sedang, rendah, sangat rendah menggunakan nilai standar deviasi dan mean.

Tabel 2 Kategorisasi data kasus konfirmasi dan meninggal COVID19

Kabupaten/Kota	Jumlah Kasus Terkonfirmasi (30/06/2021)	Jumlah Kasus Meninggal (30/06/2021)	Status Covid	Status Fatality Rate
Kab Bangkalan	3502	164	Sedang	Tinggi
Kab. Banyuwangi	7407	552	Tinggi	Tinggi
:	:	:	:	:
Kab. Blitar	6208	451	Sedang	Tinggi
Kab. Bojonegoro	2932	134	Sedang	Rendah

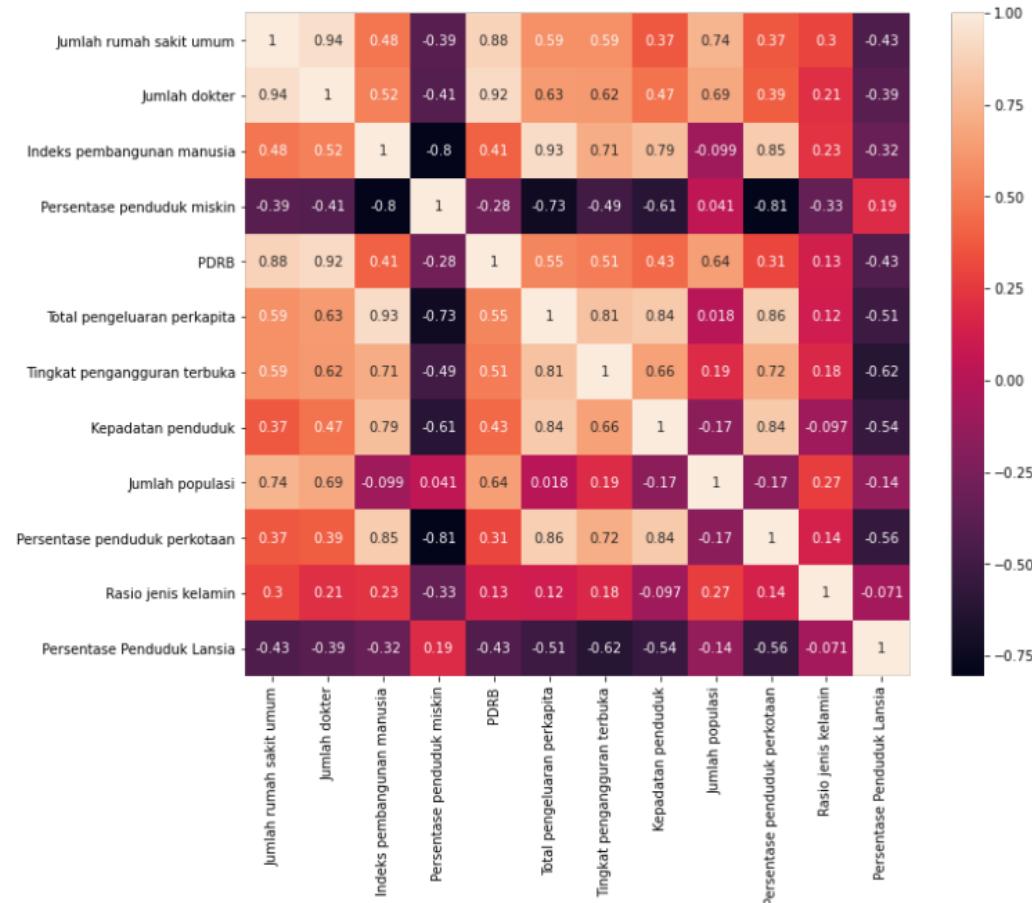
Tahapan praproses selanjutnya adalah melakukan normalisasi. Data demografi, sosial dan ekonomi di normalisasi ke dalam rentang 0-1 menggunakan normalisasi *min-max*. Hasil normalisasi dapat dilihat pada Tabel 3

Tabel 3 Normalisasi data demografi, sosial, ekonomi

Kabupaten/Kota	Jumlah Rumah Sakit Umum	...	Jumlah Dokter
Kab. Bangkalan	0.028571	...	0.030846
Kab. Banyuwangi	0.285714	...	0.235821
:	:	:	:
Kab. Blitar	0.171429	...	0.104478
Kab. Bojonegoro	0.200000	...	0.141294

- Hak Cipta Dilindungi Undang-undang  
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
 b. Pengutipan tidak mengurangi kepentingan yang wajar IPB University.  
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tahapan praproses yang terakhir adalah menghilangkan multicollinearity dengan mempertimbangkan nilai korelasi Pearson's dan Variance Inflation Factors (VIF).



Gambar 5 Korelasi Pearson's data demografi, sosial dan ekonomi

Tabel 4 Nilai VIF setelah menghilangkan beberapa atribut

Variabel	VIF Factor
Jumlah rumah sakit umum	6.9
Persentase penduduk miskin	3.3
PDRB	5.8
Total pengeluaran perkapita	9.1
Tingkat pengangguran terbuka	3.9
Kepadatan penduduk	5.1



Rasio jenis kelamin	1.6
Persentase Penduduk Lansia	2.0

## 4.2 Analisis Regresi

Data demografi, sosial dan ekonomi dimasukkan ke dalam algoritme *Multiple Linear Regression* dengan menggunakan metode *Ordinary Least Square (OLS)* untuk mendapatkan koefisien dari masing-masing variabel. Analisis regresi ini bertujuan untuk melihat variabel independen yang paling berpengaruh terhadap variabel dependen berdasarkan nilai koefisien regresi. Semakin tinggi nilai koefisien regresi, semakin tinggi juga pengaruh variabel independen terhadap variabel dependen. Variabel independen yang digunakan dalam analisis regresi ini adalah variabel PDRB, Jumlah Rumah Sakit Umum, Persentase Penduduk Lansia, Tingkat Pengangguran Terbuka, Kepadatan Penduduk, Rasio Jenis Kelamin, Persentase Penduduk Miskin dan Total Pengeluaran Perkapita. Sedangkan untuk variabel independen yang digunakan dalam analisis regresi ini adalah variabel Jumlah Kasus Konfirmasi pada tanggal 30 Juni 2021 di kabupaten/kota di Jawa Timur. Koefisien regresi dari masing-masing variabel yang sudah terpilih dapat dilihat pada Tabel 4.

Tabel 5 Koefisien Regresi Masing-Masing Variabel Terpilih

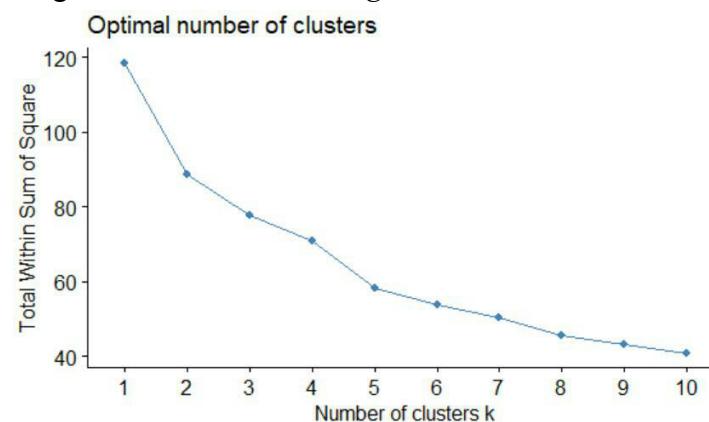
Variabel Independen	Koefisien Regresi
PDRB	0.7130
Persentase Penduduk Lansia	0.2210
Jumlah Rumah Sakit Umum	0.2074
Tingkat Pengangguran Terbuka	0.1926
Kepadatan Penduduk	0.0901
Rasio Jenis Kelamin	-0.0364
Persentase Penduduk Miskin	-0.0398
Total Pengeluaran Perkapita	-0.1389

Produk domestik bruto (PDB) berhubungan positif dengan jumlah infeksi, kematian, dan laju pertumbuhan jumlah infeksi virus. Hubungan positif yang kuat

antara COVID-19 dan produk domestik bruto menunjukkan bahwa pandemi dapat sangat mempengaruhi wilayah dengan ekonomi yang maju (Lenda *et al.* 2020).

### 4.3 Analisis *Clustering*

Pada tahap ini, data demografi, sosial, ekonomi dilakukan *clustering* menggunakan metode *K-Means*. Clustering dilakukan untuk mengelompokkan wilayah kabupaten dan kota di Jawa Timur berdasarkan kemiripan karakteristiknya. Berdasarkan metode elbow pada Gambar 6, nilai  $k=5$  cukup optimal untuk digunakan dalam *clustering*.



Gambar 6 Jumlah Cluster yang Optimal

Dari hasil *clustering* terdapat 5 cluster dengan jumlah anggota *cluster* 0 yaitu 17, *cluster* 1 yaitu 2, *cluster* 2 yaitu 11, *cluster* 3 yaitu 8 dan *cluster* 4 yaitu 2. Data hasil *cluster* digabungkan dengan data kategorisasi kasus konfirmasi dan meninggal COVID19 untuk melihat bagaimana status konfirmasi dan fatality rate COVID19 di *cluster* tersebut seperti yang terdapat pada Tabel 6

Tabel 6 *Cluster*, Wilayah, Status Covid dan Status Fatality Rate

<i>Cluster</i>	Wilayah	Status Covid	Status Fatality Rate
<i>Cluster</i> 0	Kab. Banyuwangi	Tinggi	Tinggi
	Kab. Blitar	Sedang	Tinggi
	Kab. Bojonegoro	Sedang	Rendah
	Kab. Gresik	Sedang	Sedang
	Kab. Jombang	Sedang	Sangat Tinggi
	Kab Kediri	Sedang	Tinggi
	Kab. Lamongan	Sedang	Sedang
	Kab. Lumajang	Sedang	Tinggi
	Kab Mojokerto	Sedang	Sangat Rendah
	Kab. Nganjuk	Sedang	Tinggi
	Kab. Pasuruan	Sedang	Sedang



- Hak Cipta Dilindungi Undang-undang  
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

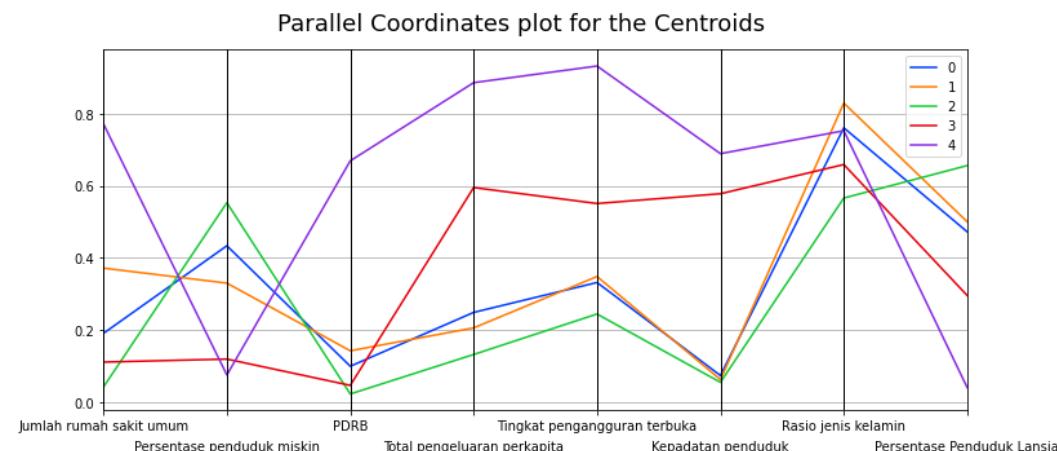
	Kab. Probolinggo	Sedang	Sedang
	Kab. Sumenep	Sedang	Sedang
	Kab. Tuban	Sedang	Sangat Tinggi
	Kab. Tulungagung	Sedang	Sangat Rendah
<i>Cluster 1</i>	Kab. Jember	Tinggi	Sedang
	Kab. Malang	Sedang	Sedang
<i>Cluster 2</i>	Kab. Bangkalan	Sedang	Tinggi
	Kab. Bondowoso	Sedang	Sedang
	Kab. Madiun	Sedang	Sedang
	Kab. Magetan	Sedang	Sedang
<i>Cluster 3</i>	Kota Batu	Rendah	Tinggi
	Kota Blitar	Sedang	Rendah
	Kota Kediri	Rendah	Tinggi
	Kota Madiun	Sedang	Sedang
	Kota Malang	Tinggi	Tinggi
	Kota Mojokerto	Sedang	Sedang
	Kota Pasuruan	Rendah	Sangat Tinggi
	Kota Probolinggo	Rendah	Sedang
<i>Cluster 4</i>	Kab. Sidoarjo	Sangat Tinggi	Rendah
	Kota Surabaya	Sangat Tinggi	Rendah

Nilai rata-rata setiap atribut pada masing-masing *cluster* dihitung untuk mengkategorikan *cluster* berdasarkan karakteristiknya. Nilai rata-rata setiap *cluster* pada tiap atribut dapat dilihat pada Tabel 7 dan Grafik pada Gambar 7 juga menampilkan visualisasi perbandingan nilai rata-rata dari data sosial, ekonomi dan demografi.

Tabel 7 Nilai rata-rata atribut tiap *cluster*

Atribut	<i>Cluster 0</i>	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>	<i>Cluster 4</i>
Jumlah rumah sakit umum	0.190476	0.371429	0.041558	0.110714	0.771429
Persentase penduduk miskin	0.433422	0.329804	0.552962	0.119044	0.074907
PDRB	0.099141	0.142003	0.022202	0.046057	0.669055
Total pengeluaran perkapita	0.248550	0.205535	0.131780	0.594967	0.885952

Tingkat pengangguran terbuka	0.331492	0.348101	0.244063	0.550633	0.932106
Kepadatan penduduk	0.073354	0.062113	0.054522	0.577783	0.688994
Rasio jenis kelamin	0.760748	0.829324	0.565865	0.658792	0.752577
Persentase Penduduk Lansia	0.472112	0.500000	0.655941	0.295483	0.040017



Gambar 7 Grafik perbandingan rataan per atribut setiap *cluster*

### Cluster Empat

Berdasarkan Gambar 7, *cluster* 4 merupakan *cluster* yang memiliki jumlah kasus konfirmasi COVID-19 yang sangat tinggi. Tingginya kasus COVID-19 di *cluster* ini berkaitan dengan jumlah rumah sakit umum yang sangat tinggi yang disebabkan oleh kepadatan penduduknya yang tinggi. Namun, tingkat kematian pada *cluster* ini cenderung rendah, hal ini menunjukkan bahwa penduduk pada *cluster* ini mayoritas merupakan penduduk usia kerja yang memiliki resiko kematian COVID-19 yang rendah. *Cluster* ini merupakan wilayah dengan perekonomian yang maju, hal ini ditunjukan dari nilai pdrb yang sangat tinggi juga total pengeluaran perkapita yang tinggi. *Cluster* ini juga merupakan daerah perdagangan dan industri yang mana berdiri juga perusahaan besar seperti Sampoerna, Unilever, Bogasari, dan sebagainya.

### Cluster Dua

Pada *cluster* 2, *cluster* ini merupakan *cluster* dengan Wilayah dengan status covid yang cenderung rendah dan sedang. Perekonomian pada *cluster* ini masih kurang maju dilihat dari persentase penduduk miskinnya yang tinggi serta



pdrb dan total pengeluaran perkapita yang rendah. Wilayah yang berada pada *cluster* ini berada pada sektor pertanian dan perdagangan paling dominan. Rasio jenis kelamin yang rendah pada *cluster* ini juga menggambarkan risiko infeksi COVID-19 pada *cluster* ini. Berdasarkan penelitian Dörre dan Doblhammer (2020) bahwa rasio jenis kelamin menunjukkan risiko infeksi yang lebih tinggi diantara perempuan dibandingkan laki-laki pada usia kerja. Rendahnya rasio jenis kelamin di wilayah *cluster* ini menunjukkan risiko infeksi COVID-19 pada perempuan yang tinggal di wilayah ini cenderung rendah. Kemudian, persentase penduduk lansia di wilayah *cluster* ini juga cenderung tinggi, hal ini menunjukkan bahwa mayoritas penduduk merupakan penduduk dengan usia lanjut.

### **Cluster Nol**

Pada *cluster* 0, sebagian besar kasus COVID-19 pada wilayah ini berstatus sedang. Kabupaten Banyuwangi merupakan satu-satunya wilayah di *cluster* ini yang memiliki status COVID-19 tinggi. Wilayah yang berada pada *cluster* ini berada pada sektor pertanian dan perdagangan paling dominan. *Cluster* ini juga memiliki kepadatan penduduk yang rendah, kepadatan penduduk yang rendah mengakibatkan ketersediaannya lapangan pekerjaan yang membuat tingkat pengangguran terbuka menjadi rendah pada *cluster* ini.

### **Cluster Satu**

Pada *cluster* 1, status COVID-19 pada wilayah *cluster* ini adalah tinggi dan sedang. *Cluster* ini hanya terdiri dari dua wilayah yaitu Kabupaten Jember dan Kabupaten Malang. Jika dilihat dari total pengeluaran perkapitanya yang rendah dapat dilihat bahwa daya beli masyarakat pada *cluster* ini dinilai cukup rendah. Kemudian, rasio jenis kelamin yang tinggi pada *cluster* ini menunjukkan bahwa risiko infeksi COVID-19 pada perempuan yang tinggal di wilayah *cluster* ini cenderung tinggi.

### **Cluster Tiga**

Pada *cluster* 3, merupakan *cluster* dengan seluruh wilayah kota di Jawa Timur, kecuali Kota Surabaya. Kota malang merupakan satu satunya kota yang memiliki status COVID-19 dan tingkat kematian yang tinggi pada *cluster* ini. Berdasarkan penelitian Lenda *et al.* (2020) bahwa hubungan positif yang kuat antara COVID-19 dan produk domestik bruto menunjukkan bahwa pandemi dapat sangat mempengaruhi wilayah dengan ekonomi yang maju. PDRB yang rendah pada wilayah *cluster* ini menunjukkan bahwa perekonomian wilayah ini dinilai masih kurang maju dan pandemi tidak cukup mempengaruhi wilayah *cluster* ini. Namun disamping rendahnya PDRB di wilayah *cluster* ini, kontributor utama pendapatan di wilayah *cluster* ini berada pada sektor perdagangan dan pariwisata. Selain itu, rendahnya rasio jenis kelamin di wilayah *cluster* ini juga menunjukkan



bahwa risiko infeksi COVID-19 pada perempuan yang tinggal di wilayah *cluster* ini cenderung rendah.

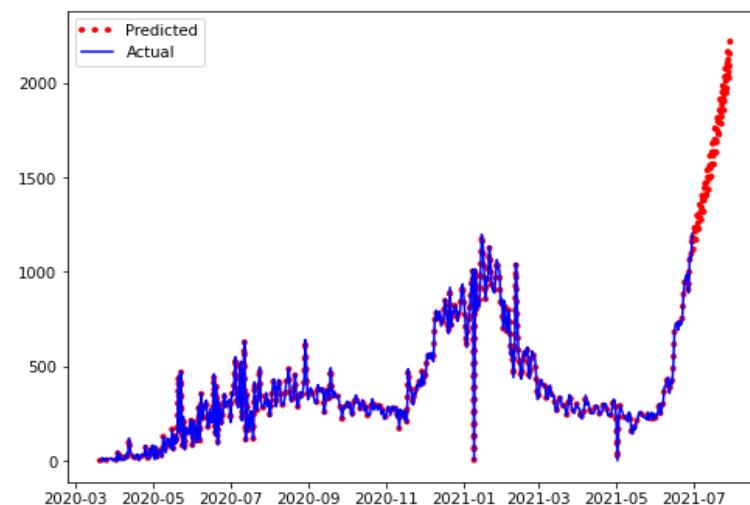
#### 4.4 Pemodelan Prediksi Menggunakan LSTM

Data yang digunakan untuk membangun model prediksi adalah data kasus konfirmasi COVID19. Data kemudian dibagi menjadi 70% data latih dan 30% data uji. Model dibuat menggunakan metode LSTM yang ada di paket Keras pada Python. Arsitektur LSTM untuk model prediksi kasus COVID19 pada penelitian ini ditunjukkan pada Tabel 8. Penelitian ini menggunakan 1 LSTM *layer* dan 1 Dense *layer*. Fungsi aktivasi yang digunakan dalam LSTM *layer* adalah *relu*. Untuk *optimizer* yang digunakan dalam arsitektur LSTM adalah *Adaptive Moment Estimation* (Adam), *epoch* yang digunakan adalah sebanyak 100 *epoch* dan batch size yang digunakan sebanyak 1.

Tabel 8 Struktur LSTM

Karakteristik	Spesifikasi
Arsitektur	1 LSTM layer dengan 4 unit 1 Dense layer
Fungsi aktivasi	relu
Optimizer	Adam
Epoch	100
Batch size	1

Berdasarkan Gambar 8, model LSTM sudah baik dalam memprediksi kasus COVID-19 di Jawa Timur. Hal ini terlihat dari output data prediksi yang sudah mendekati output data aktual. Berdasarkan hasil prediksi selama 30 hari kedepan dari tanggal 1 Juli 2021 sampai 30 Juli 2021, ini menunjukan bahwa kasus COVID-19 di wilayah Jawa Timur masih terus meningkat secara drastis dan belum ada penurunan.



Gambar 8 Plot perbandingan nilai aktual dan prediksi kasus konfirmasi COVID19

#### 4.5 Evaluasi Model

Pada tahap ini dilakukan evaluasi model untuk melihat tingkat kebaikan dari model yang dibentuk. Evaluasi dilakukan dengan menghitung nilai *Root Mean Squared Error* (RMSE). Pada pemodelan yang sudah dilakukan, didapatkan nilai train RMSE sebesar 127.96 dan test RMSE sebesar 79.83. Semakin kecil nilai RMSE yang dihasilkan, menunjukkan bahwa variasi nilai yang dihasilkan suatu model mendekati variasi nilai aktualnya.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## V SIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan analisis deskriptif yang dilakukan dalam penelitian ini, ditemukan bahwa terdapat hubungan antara faktor demografi, sosial dan ekonomi dengan jumlah kasus konfirmasi COVID-19 di Provinsi Jawa Timur. Faktor yang memiliki pengaruh signifikan yaitu nilai PDRB. Ini terbukti dari *cluster* yang memiliki jumlah kasus konfirmasi COVID-19 sangat tinggi juga memiliki nilai PDRB yang sangat tinggi. Namun, ada daerah dengan kasus COVID-19 yang tinggi, tetapi tidak bisa dijelaskan dengan nilai dari variabel PDRB. Sehingga variabel yang digunakan pada penelitian ini, tidak bisa menangkap keseluruhan kondisi COVID-19 di Jawa Timur. Jumlah positif COVID-19 terbanyak di Jawa Timur ditemukan di daerah pusat perekonomian seperti Kota Surabaya dan Kabupaten Sidoarjo serta daerah-daerah yang terhubung dengan wilayah pusat perekonomian ini. Pada analisis prediktif, teknik LSTM yang digunakan untuk memprediksi jumlah penambahan kasus konfirmasi COVID-19 dinilai sudah cukup baik karena sudah mengikuti pola dari data aktualnya. Diperkirakan penambahan jumlah positif COVID-19 masih akan terus naik drastis hingga akhir Juli 2021. Jika dikaitkan dengan analisis *cluster*, peningkatan kasus paling banyak kemungkinan terjadi di daerah pusat perekonomian serta daerah-daerah yang terhubung dengan wilayah ini.

### 5.2 Saran

Analisis deskriptif pada penelitian ini hanya menggunakan 12 atribut sosial, ekonomi dan demografi. Sedangkan dalam penyebaran kasus COVID-19 ada banyak faktor lainnya yang mempengaruhi seperti faktor cuaca, mobilitas masyarakat, serta kebijakan daerah setempat. Oleh karena itu, pada penelitian berikutnya faktor-faktor tersebut bisa dipertimbangkan dan diamati lebih lanjut. Kemudian pada analisis prediktif pada penelitian ini hanya menggunakan variabel penambahan jumlah positif COVID-19, belum mempertimbangkan adanya pengaruh dari variabel lainnya seperti variabel meninggal, sembuh atau yang lainnya. Sehingga, pada penelitian berikutnya variabel lainnya bisa dimasukkan sebagai input pada model, untuk mendapatkan informasi yang lebih detail. Selain itu analisis prediktif dapat dilakukan menggunakan metode lain untuk dilakukan perbandingan



- Achim Dorre, Gabriele Doblhammer. 2020. The Effect of Gender on Covid-19 Infections and Mortality in Germany: Insights From Age- and Sex-Specific Modelling of Contact Rates, Infections, and Deaths. DOI: 10.1101/2020.10.06.20207951.
- Adhinata FD, Rakhmadani DP. 2021. Prediction of Covid-19 Daily Case in Indonesia Using Long Short Term Memory Method. DOI: 10.34148/teknika.v10i1.328.
- Aiken LS, West SG, Pitts, Baraldi AN, Wurpts IC. 2012. Multiple Linear Regression. DOI: 10.1002/9781118133880.hop202018.
- Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F. Koczy ARV, Reuter U, Rabczuk T, Atkinson PM. 2020. COVID-19 Outbreak Prediction with Machine Learning. DOI: 10.1101/2020.04.17.20070094v1.
- Bavel VJ. 2013. The world population explosion: causes, backgrounds and projections for the future. *Facts Views Vis Obgyn.* 5(4):281-291.
- Bodapati S, Bandarupally H, Trupthi H. 2020. COVID-19 Time Series Forecasting of Daily Cases, Deaths Caused and Recovered Cases using Long Short Term Memory Networks. DOI: 10.1109/ICCCA49541.2020.9250863.
- [BPS] Badan Pusat Statistik. 2019. Penduduk, Laju Pertumbuhan Penduduk, Distribusi Persentase Penduduk, Kepadatan Penduduk, dan Rasio Jenis Kelamin Penduduk Menurut Provinsi [internet]. [diunduh 2020 Nov 23]. Tersedia pada: [https://www.bps.go.id/indikator/indikator/view\\_data\\_pub/0000/api\\_pub/50/da\\_03/1](https://www.bps.go.id/indikator/indikator/view_data_pub/0000/api_pub/50/da_03/1).
- Cao X. 2020. COVID-19: immunopathology and its implications for therapy. *Nat Rev Immunol.* 20:269–270. DOI: 10.1038/s41577-020-0308-3.
- Chi-Wei Su, Ke Dai, Sana Ullah, Zubaria Andlib. 2021. COVID-19 pandemic and unemployment dynamics in European economies. DOI: 10.1080/1331677X.2021.1912627.
- Chae S, Kwon S, Lee D. 2018. Predicting Infectious Disease Using Deep Learning and Big Data. *Int. J. Environ. Res. Public Health.* 15(8). DOI:10.3390/ijerph15081596
- Cheng SC, Chang YC, Chiang YLF, Chien YC, Cheng M, Yang Ch, Huang CH, Hsu YN. 2020. First case of Coronavirus Disease 2019 (COVID-19) pneumonia in Taiwan. *Journal of the Formosan Medical Association.* 119(3):745-751. DOI:10.1016/j.jfma.2020.02.007.
- Grassly NC, Fraser C. 2008. Mathematical models of infectious disease transmission. *Nat Rev Microbiol.* 6:477-487. DOI:10.1038/nrmicro1845.
- Guernier V, Hochberg ME, Guegan JF. 2004. Ecology Drives the Worldwide Distribution of Human Diseases. *PLoS Biol.* 2(6): e141. DOI: 10.1371/journal.pbio.0020141.
- Heesterbeek H, Anderson RM, Andreasen V, Bansal S, Angelis DD, Dye C, Eames KTD, Edmunds WJ, Frost SDW, Funk S, et al. 2015. Modeling infectious disease dynamics in the complex landscape of global health. *Science.* 347(6227). DOI: 10.1126/science.aaa4339.

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



- Hochreiter S, Schmidhuber J. 1997. Long Short-Term Memory. *Neural Computation*. 9:1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- Joanne Brooke, Debra Jackson. 2020. Older people and COVID-19: Isolation, risk and ageism. DOI: 10.1111/jocn.15274.
- Junior Hutagalung, Ni Luh WSRG, Gita WB, Wayan GSP, Anjar Wantu, Pawer DP. 2020. COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm. DOI: 10.1088/1742-6596/1783/1/012027.
- Kalchbrenner N, Danihelka I, Graves A. 2016. Grid Long Short-Term Memory. *arXiv.org*:1507.01526.
- Karlic B, Olgac AV. 2016. Performance analysis of various activation functions on generalized MLP architectures of neural network. *International Journal of Artificial Intelligent and Expert System (IJAE)*. 1(4):112.
- Karpathy A, Johnson J, Fei-Fei L. 2015. Visualizing and understanding recurrent networks. *arXiv*:1506.02078.
- [KPCPEN] Komite Penanganan COVID-19 dan Pemulihan Ekonomi Nasional. 2020. Peta Sebaran [internet]. [diunduh 2020 Nov 18]. Tersedia pada: <https://covid19.go.id/peta-sebaran>.
- Liping Y, Yuntao P, Yishan W. 2009. Research on Data Normalization methods in Multi-attribute Evaluation. DOI: 10.1109/CISE.2009.5362721.
- Liu L, Han M, Zhou Y, Wang Y. 2018. LSTM Recurrent Neural Networks for Influenza Trends Prediction. DOI: 10.1007/978-3-319-94968-0\_25.
- Mahmudan A. 2020. Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering. DOI: 10.20956/jmsk.v%vi%i.10727.
- Marinescu AI. 2019. Bach 2.0 - generating classical music using recurrent neural networks. *Procedia Computer Science*. 159:117-124. DOI:10.1016/j.procs.2019.09.166.
- Morse SS. 1995. Factors in the emergence of infectious diseases. *Emerging Infectious Diseases*. 1(1):7-15. DOI:10.3201/eid0101.950102.
- Piotr Skorka, Beata Grzywacz, Dawid Moron, Magdalena Lenda. 2020. The macroecology of the COVID-19 pandemic in the Anthropocene. DOI: 10.1371/journal.pone.0236856.
- Pemerintah Provinsi Jawa Timur. 2019. Data dinamis perekonomian Jawa Timur [internet]. [diunduh 2020 Nov 22]. Tersedia pada: <https://ro-ekonomi.jatimprov.go.id/storage/pdf/data%20dinamis/HN6bXWnyb4SnVGVNSLb3RRqKSvGUyz2rViembbw.pdf>.
- Ragab D, Eldin HS, Taeimah M, Khattab R, Salem R. 2020. The COVID-19 Cytokine Storm: What We Know So Far. *Front. Immunol.* 11(1446). DOI:10.3389/fimmu.2020.01446.
- Rodgers JL, Nicewander WA. 1988. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*. 42(1):59-66. DOI:10.1080/00031305.1988.10475524.
- Stein ML, Steenbergen JEV, Buskens V, Heijden PGMV, Chanyasanha C, Tipayangmongkholgul M, Thorson AE, Bengtsson L, Lu X, Kretzschmar MEE. 2014. Comparison of Contact Patterns Relevant for Transmission of Respiratory Pathogens in Thailand and the Netherlands Using

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

- Respondent-Driven Sampling. *PLoS ONE*. 9(11): e113711. DOI: 10.1371/journal.pone.0113711.
- Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, Chen H, Wang D, Liu N, Chen G, et al. 2020. Characteristics of COVID-19 infection in Beijing. *Journal of Infection*. 80(4):401-406. DOI: 10.1016/j.jinf.2020.02.018.
- Turista DDR, Islamy A, Kharisma VD, Ansori ANM. 2020. Distribution of COVID-19 and Phylogenetic Tree Construction of SARS-CoV-2 in Indonesia. *Journal of Pure and Applied Microbiology*. 14(1):1035-1042. DOI:10.22207/JPAM.14.SPL1.42.
- [WHO] World Health Organization. 2020. Weekly epidemiological update - 3 November 2020 [internet]. [diunduh 2020 Nov 1]. Tersedia pada: <https://www.who.int/publications/m/item/weekly-epidemiological-update--3-november-2020>.
- [WHO] World Health Organization. 2020. Coronavirus disease (COVID-19): Vaccines [internet]. [diunduh 2020 Nov 30]. Tersedia pada: [https://www.who.int/news-room/q-a-detail/coronavirus-disease-\(covid-19\)-vaccines](https://www.who.int/news-room/q-a-detail/coronavirus-disease-(covid-19)-vaccines).
- Worldometer. 2020. COVID-19 Coronavirus Pandemic [internet]. [diunduh 2020 Nov 15]. Tersedia pada: <https://www.worldometers.info/coronavirus>.
- Zhang Z, Fu R, Li L. 2016. Using LSTM and GRU Neural Networks Methods for Traffic Flow Prediction. DOI: 10.1109/YAC.2016.7804912.



## RIWAYAT HIDUP

Penulis dilahirkan di Kota Jakarta pada tanggal 28 November 1999. Penulis dilahirkan sebagai anak pertama dari lima bersaudara oleh Bapak Avries Praputro Wulandoro dan Ibu Swarini Wigati. Pada tahun 2014, penulis menempuh pendidikan di SMA Negeri 101 Jakarta yang kemudian melanjutkan studi pada Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, IPB University pada tahun 2017 melalui jalur SNMPTN.

Penulis juga mengikuti beberapa kegiatan di kampus seperti pada kegiatan yang dilaksanakan oleh Himalkom yaitu IT Today 2018 dan 2019 sebagai anggota dan ketua divisi entrepreneur dan konsumsi, kegiatan fakultas yaitu Pesta Sains Nasional (PSN) 2018 sebagai anggota divisi publikasi, dekorasi dan desain (PDD) serta kegiatan IPB Art Contest (IAC) 2020 sebagai bendahara divisi publikasi, multimedia dan desain. Selain itu, penulis juga menjadi asisten praktikum beberapa mata kuliah yaitu Pengantar Komputer, Pengantar Matematika Komputasi, Interaksi Manusia Komputer dan Data Mining.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak mengurangi kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.