

# Prediksi Penyakit Daun Pisang Menggunakan Metode LSTM (Long Short-Term Memory)

Alfian Kafilah Ba'its<sup>1</sup>, Andika Setiawan<sup>2</sup>, Winda Yulita<sup>3</sup>, Harmiansyah<sup>4</sup>, Amalia Listiani<sup>5</sup>, Meida Cahyo Untoro<sup>6</sup>, Nike Dwi Grevika Drantantiyas<sup>7</sup>, Amir Faisal<sup>8</sup>, Leslie Anggraini<sup>9</sup>, Andre Febrianto<sup>10</sup>, Mohamad Meazza Aprilianda<sup>11</sup>, Mhd. Kadar Fitrawan<sup>12</sup>, Radhinka Bagaskara<sup>\*13</sup>

<sup>1,2,3,6,9,10,11</sup>Teknik Informatika, Institut Teknologi Sumatera

<sup>4,13</sup>Teknik Biosistem, Institut Teknologi Sumatera

<sup>5</sup>Sains Aktuaria, Institut Teknologi Sumatera

<sup>7</sup>Teknik Fisika, Institut Teknologi Sumatera

<sup>8</sup>Teknik Biomedis, Institut Teknologi Sumatera

e-mail: <sup>1</sup>alfian.120140109@student.itera.ac.id, <sup>2</sup>andika.setiawan@if.itera.ac.id,

<sup>3</sup>winda.yulita@if.itera.ac.id, <sup>4</sup>harmiansyah@tbs.itera.ac.id, <sup>5</sup>amalia.listiani@at.itera.ac.id,

<sup>6</sup>cahyo.untoro@if.itera.ac.id, <sup>7</sup>nike.drantantiyas@tf.itera.ac.id, <sup>8</sup>amir.faisal@bm.itera.ac.id,

<sup>9</sup>leslie.anggraini@if.itera.ac.id, <sup>10</sup>andre.febrianto@if.itera.ac.id,

<sup>11</sup>mohamad.121140216@student.itera.ac.id, <sup>12</sup>mhd.120310059@student.itera.ac.id,

<sup>\*13</sup>radhinka.bagaskara@if.itera.ac.id

## Abstrak

*Dalam sektor pertanian, tanaman yang memiliki peran signifikan dalam skala global adalah pisang, yaitu buah yang mudah didapatkan, dapat tumbuh dimana saja, memiliki gizi yang tinggi, serta memiliki nilai ekonomi & budaya yang tinggi. Pisang mempunyai kontribusi yang signifikan terhadap pendapatan nasional Indonesia, terutama di Provinsi Lampung sebagai penghasil pisang nasional terbesar. Tetapi, proses produksi pisang seringkali mengalami kendala, salah satunya karena faktor serangan penyakit Black Sigatoka. Penyakit tersebut memberikan kerugian pada tanaman pisang, seperti daun yang meranggas, panen tertunda, bakal buah rontok, dan kualitas buah yang rendah, dan dapat menyebar melalui aliran udara atau percikan air hujan. Tingkat keparahan penyakit Black Sigatoka perlu diprediksi agar penyakit tersebut dapat dikontrol dan dapat dicegah sedini mungkin. Model yang digunakan untuk memprediksi permasalahan ini dalam jangka panjang adalah model Long Short-Term Memory (LSTM), salah satu jenis dari arsitektur Recurrent Neural Network (RNN), yang mempunyai kinerja yang baik dan mempunyai model yang prediktif. Aplikasi LSTM diterapkan terhadap dataset pohon pisang yang terdampak penyakit Black Sigatoka. Hasil dari model LSTM dalam melakukan prediksi penyakit Black Sigatoka menghasilkan model dengan nilai error yang kecil, dengan nilai MAE dan MAPE masing-masing sebesar 0.084 dan 5.7%.*

**Kata kunci**—Pisang, Black Sigatoka, Prediksi, LSTM

## Abstract

*In the agriculture sector, a fruit crop with a significant role in global scale is banana, due to its accessibility, ability to grow in almost everywhere, high nutrition value, and having a high economical & cultural value. Banana had contributed significantly towards Indonesia's national income, especially in the Province of Lampung as the national's biggest banana producer. While Province of Lampung's banana exports is enormous, however during the production process there are issues, one of which is Black Sigatoka disease attacks. Said disease negatively impacting banana production, with effects such as withering leaves, delayed crops, decaying fruits, and lower fruit qualities. The severity of Black Sigatoka needed to be predicted in order to control and prevent the disease as early as possible. The model for long term prediction is Long Short-Term Memory (LSTM) model, a model from Recurrent Neural Network (RNN) architecture, with a good performance and predictive model. LSTM applicated toward dataset contains banana trees that are affected by Black Sigatoka disease. The results is an LSTM model to predict Black Sigatoka disease with small errors, where the MAE and MAPE scores of 0.084 and 5.7% respectively.*

**Keywords**—Banana, Black Sigatoka, Prediction, LSTM

## 1. PENDAHULUAN

Komoditas buah pisang merupakan salah satu sektor pertanian yang memiliki peran signifikan, dikarenakan kontribusinya terhadap pendapatan nasional Indonesia terutama untuk sektor ekspor [1]. Badan Pusat Statistik mencatat bahwa Provinsi Lampung menjadi penghasil pisang terbesar setelah Provinsi Jawa Timur dan Jawa Barat dengan volume produksi mencapai 1.22 juta ton pada tahun 2022 [2]. Ekspor yang dilakukan mencapai 4.02 miliar dolar AS pada periode Januari - Oktober 2021, menunjukkan bahwa permintaan pasar ekspor buah pisang sangat besar. Akan tetapi, jumlah produksi pisang di Provinsi Lampung turun dari 1.20 juta ton pada tahun sebelumnya menjadi sebesar 1.12 juta ton pada tahun 2021 [2]. Penurunan produksi ini disebabkan oleh penyakit Black Sigatoka.

Penyakit Black Sigatoka merupakan penyakit bercak daun pada tanaman pisang yang disebabkan oleh jamur *Mycosphaerella Fijensis* [3]. Terdapat tiga klasifikasi pada penyakit Black Sigatoka, antara lain Early Streak, Early Spot, dan BLS Mature. Dampak yang terjadi akibat penyakit ini, seperti daun dan buah yang lebih sedikit serta lebih kecil, daun yang meranggas, panen tertunda, bakal buah rontok, dan kualitas buah yang rendah. Black Sigatoka merupakan sebuah penyakit yang sulit untuk meramalkan kapan sebuah pohon pisang terjangkit penyakit beserta tingkat keparahannya pada waktu yang akan datang. Hal ini yang sangat merugikan dalam sektor industri pisang. Penanganan seperti prediksi tingkat keparahan penyakit Black Sigatoka perlu dilakukan agar penyakit tersebut dapat dikontrol dan dapat dicegah sedini mungkin, serta mencegah kerugian ekonomi. Terdapat metode prediksi seperti pendekatan statistika, pendekatan machine learning, atau deep learning. Metode prediksi tersebut dapat diterapkan dalam industri pertanian pisang, untuk mengambil tindakan yang tepat untuk mengurangi penyebaran penyakit Black Sigatoka pada tanaman pisang.

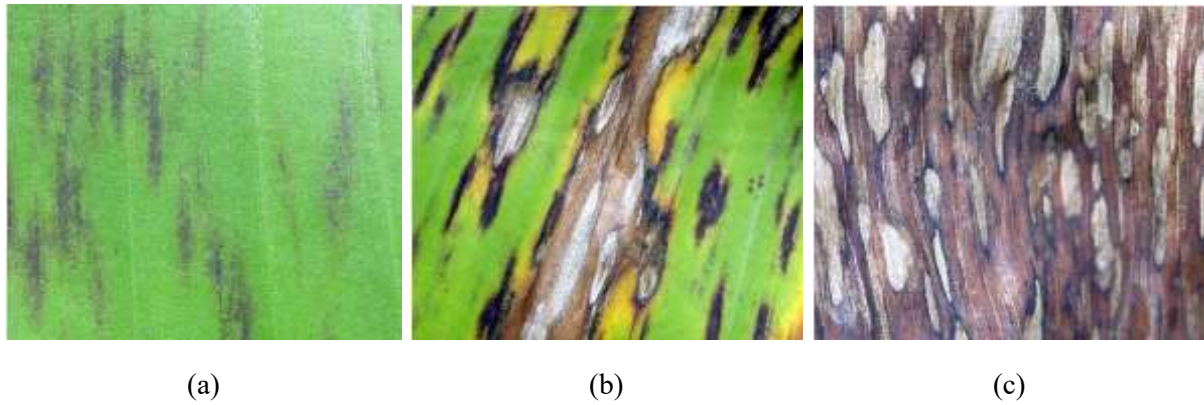
Model prediksi (forecasting) merupakan suatu kegiatan untuk melakukan prediksi terhadap suatu objek secara kuantitatif untuk mengetahui apa yang akan terjadi pada masa yang akan datang dengan mengolah data masa lalu [4]. Prediksi dilakukan dengan memanfaatkan informasi pada masa lalu dan analisis time series untuk menimbang kegiatan di masa yang akan datang. Terdapat berbagai pemodelan time series yang dapat digunakan seperti ARMA, ARIMA, Prophet, ataupun deep learning seperti LSTM [5]. Metode ARMA, ARIMA ataupun Prophet tidak mempunyai kinerja yang baik dalam mengolah data historikal yang panjang. Salah satu model yang sering digunakan untuk masalah peramalan jangka panjang adalah model Long Short-Term Memory (LSTM).

Long Short-Term Memory (LSTM) merupakan salah satu jenis dari arsitektur Recurrent Neural Network (RNN) yang mampu mengatasi keterbatasan memori jangka pendek [6]. LSTM menjadi pilihan terbaik dibandingkan metode lain, karena metode ini dapat melakukan pemodelan dalam mempelajari hubungan jangka panjang antar data. Hal ini dapat ditunjukkan dari penelitian sebelumnya yang membandingkan metode LSTM dengan metode lainnya, didapati bahwa metode LSTM mempunyai kinerja yang baik dan mempunyai model yang prediktif [7]. Dari penelitian yang dilakukan oleh Diah Aisyah (2023), didapati hasil pengujian dengan parameter nilai RMSE sebesar 0.12, MAE sebesar 0.08, dan R2 sebesar 0.53 [7]. Penelitian Ayu Ahadi Ningrum (2021), menunjukkan bahwa algoritma Deep Learning-LSTM mempunyai kinerja yang baik dengan nilai RMSE sebesar 0.0004 dan nilai Squared Correlation sebesar 0.9690 [8].

Penggunaan algoritma Long Short-Term Memory dapat digunakan untuk prediksi jangka panjang karena algoritma ini mempunyai memori sel yang mampu menyimpan informasi dalam jangka waktu yang panjang [9]. Berdasarkan permasalahan dan penelitian sebelumnya, maka dilakukan sebuah penelitian dengan menggunakan metode LSTM dan ke arah objek yang berbeda, dengan harapan pemodelan peramalan ini dapat berkontribusi dalam upaya pencegahan dan perlindungan tanaman pisang dari penyakit Black Sigatoka.

## 2. METODE PENELITIAN

Penelitian dilaksanakan dengan melakukan pengambilan data yang dimulai dari tanggal 31 Maret 2019 hingga 2 Juli 2019. Data yang didapatkan berupa data time series dalam mingguan tentang tingkat keparahan penyakit *Black Sigatoka* pada masing-masing daun. Data tersebut terdiri dari 3 kolom tentang tingkat keparahan penyakit, yaitu Early Streak, Early Spot, dan BLS Mature.

Gambar 1. Citra tingkat penyakit *Black Sigatoka*

Metode yang digunakan dalam penelitian ini adalah Long Short-Term Memory (LSTM), yang merupakan salah satu jenis arsitektur Recurrent Neural Network (RNN). LSTM dipilih karena kemampuannya dalam mengatasi keterbatasan memori jangka pendek dan mempelajari hubungan jangka panjang antar data. Arsitektur LSTM menggunakan struktur yang berdasarkan proses memori jangka pendek untuk membuat memori jangka panjang, sehingga mampu mengatasi permasalahan vanishing gradient atau exploding gradient.

Pengujian yang dilakukan dalam penelitian ini dibagi menjadi beberapa bagian. Pertama, mekanisme penilaian dan validasi hasil penelitian dilakukan oleh pihak terkait dengan cara melakukan pengujian di lapangan secara langsung. Pengujian ini melibatkan identifikasi dan penyesuaian terhadap hasil peramalan yang telah dilakukan. Validasi hasil penelitian ini memberikan dasar untuk memberikan saran-saran dalam pengembangan model prediksi baru berdasarkan hasil yang telah didapatkan. Selain itu, evaluasi model dilakukan dengan menguji kinerja model pada data yang tidak dikenal, seperti data testing yang belum digunakan dalam pelatihan. Evaluasi ini penting untuk memastikan performa model yang sesuai dan akurat, dengan menggunakan metrik seperti Mean Absolute Error (MAE) dan Mean Absolute Percent Error (MAPE) untuk mengukur ketepatan prediksi.

Tabel 1 Sampel dataset awal

Tanggal	Blok	Plot	Early Streak	Severe	BLS Mature
3/31/2019	009B	1	xxx	xxx	xxx
3/31/2019	009B	2	xxx	xxx	xxx
...	...	...	...	...	...
7/2/2019	009B	9	xxx	xxx	xxx

Data yang terlihat pada Tabel 1 terdiri dari 3 kolom, yaitu Early Streak, Early Spot, dan BLS Mature. 3 jenis kolom tersebut juga mempunyai sub-kolom yang berisikan baris pohon, yaitu baris 1, baris 3, baris 5, baris 7, baris 9, dan baris 11.

Dataset awal yang telah diperoleh selanjutnya akan diproses dengan tahapan preprocessing data yang mencakup data cleaning, data transformation, dan data augmentation. Data cleaning merupakan proses yang berfungsi untuk membersihkan data dengan menghapus data yang tidak digunakan serta mempertahankan data yang diperlukan. Pada tahap ini, atribut yang tersisa adalah tanggal, blok, early streak, early spot, dan BLS Mature. Hasil dari proses ini hanya menghilangkan kolom yang tidak diperlukan, namun jumlah data tetap sama. Data transformation adalah proses yang digunakan untuk mengubah format, struktur, atau nilai data menjadi format yang dapat dianalisis untuk mendukung tujuan tertentu. Hasil dari proses data transformation bertujuan untuk mempermudah analisis dan prediksi pada data. Data augmentation adalah proses yang digunakan untuk meningkatkan jumlah data yang berasal dari data asli, sehingga data yang digunakan untuk pelatihan menjadi lebih banyak dan mampu meningkatkan akurasi model machine learning yang dibuat. Pada kesempatan ini, data augmentation dilakukan dengan menambahkan data dari gejala penyakit pada hari sebelumnya.

Hasil dari data yang telah diproses melalui preprocessing selanjutnya akan melalui proses label encoding. Label encoding merupakan salah satu teknik yang digunakan untuk mengubah data kategorikal menjadi data numerik. Hal ini dilakukan karena komputer tidak dapat memproses data bertipe kategori. Proses ini menghasilkan data numerik yang memiliki hubungan ordinal dengan kondisi masing-masing daun dan menunjukkan adanya urutan tertentu, seperti yang terlihat pada Tabel 2.

Tabel 2 Sampel hasil label encoding

Tanggal	X_0	...	X_44	X_45	X_46	X_47	X_48	X_49	X_50
3/31/2019	0	...	xxx	xxx	xxx	xxx	xxx	xxx	xxx
4/1/2019	0	...	xxx	xxx	xxx	xxx	xxx	xxx	xxx
4/2/2019	0	...	xxx	xxx	xxx	xxx	xxx	xxx	xxx
...	...	...	...	...	...	...	...	...	...
7/2/2019	0	...	xxx	xxx	xxx	xxx	xxx	xxx	xxx

Hasil data label encoding akan dibagi menjadi dua buah dataset: dataset training & dataset testing. Data training merupakan bagian dari dataset yang digunakan untuk proses pelatihan pada model machine learning. Kemudian, data testing itu juga bagian dari dataset yang digunakan untuk melakukan evaluasi terhadap kinerja model machine learning yang telah dibuat. Tujuan dari splitting data adalah untuk mendapatkan model dengan akurasi yang baik dan memastikan bahwa kinerja dari model tersebut dapat bekerja dengan baik terhadap data yang belum pernah dilihat sebelumnya. Rasio pembagian dataset antara training & testing adalah 80:20.

Hasil dari proses label encoding akan dibagi menjadi dua dataset: dataset training dan dataset testing. Dataset training adalah bagian dari dataset yang digunakan untuk melatih model machine learning. Sementara itu, dataset testing merupakan bagian dari dataset yang digunakan untuk mengevaluasi kinerja model machine learning yang telah dibuat. Tujuan dari pembagian data ini adalah untuk memperoleh model dengan akurasi yang baik dan memastikan bahwa model tersebut mampu berfungsi dengan baik pada data yang belum pernah dilihat sebelumnya. Rasio pembagian antara dataset pelatihan dan pengujian adalah 80:20.

Tahap berikutnya adalah pengembangan model LSTM untuk mencapai akurasi terbaik. Proses pembuatan model prediksi dimulai dengan inialisasi variabel model Sequential. Model Sequential ini terdiri dari lapisan-lapisan yang terhubung secara berurutan, di mana output dari lapisan sebelumnya menjadi input untuk lapisan berikutnya. Semakin banyak neuron tersembunyi yang digunakan pada sebuah lapisan, semakin kompleks pola yang harus dipelajari oleh model. Hal ini dapat menyebabkan overfitting saat proses pelatihan. Lapisan pertama dimulai dengan input layer. Pada lapisan kedua, digunakan fungsi aktivasi ReLU. Kemudian, pada lapisan ketiga, terdapat Dense layer yang menggunakan fungsi aktivasi linear. Dense layer adalah lapisan yang sepenuhnya terhubung, di mana setiap neuron di lapisan ini terhubung dengan setiap neuron di lapisan sebelumnya. Model akan menyesuaikan bobot dan bias pada Dense layer ini selama pelatihan, sehingga terbentuk model yang akurat dan optimal untuk memprediksi penyakit *Black Sigatoka*.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pengumpulan Data

Dataset yang diambil berupa data *time series* yang hanya berfokus pada perpindahan kondisi penyakit *Black Sigatoka* yang menginfeksi daun tertentu dari kondisi awal (belum tumbuh) hingga kondisi BLS *Mature* pada tiap minggunya. Kondisi tersebut bersifat ordinal dan menunjukkan adanya urutan. Tabel 3 merupakan ringkasan dataset awal yang didapatkan.

Tabel 3 Dataset awal

Tanggal	Blok	Plot	Early Streak						Severe						BLS Mature						...
			1	3	5	7	9	11	1	3	5	7	9	11	1	3	5	7	9	11	...
3/31/2019	009B	1	6	5	6	0	0	0	0	0	0	0	0	0	7	7	7	0	0	0	...
3/31/2019	009B	2	5	5	5	5	5	5	0	6	0	7	0	7	7	7	6	9	9	8	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
7/2/2019	009B	8	5	6	5	4	5	4	6	0	7	6	6	5	9	10	12	9	10	8	...
7/2/2019	009B	9	5	6	4	0	0	0	6	0	5	0	0	0	12	11	10	0	0	0	...

### 3.2 Preprocessing Dataset

Dataset yang telah dikumpulkan, selanjutnya dilakukan *preprocessing* dataset dengan beberapa tahapan, antara lain:

#### 1. Data Cleaning

*Data cleaning* bertujuan untuk memastikan kebersihan dan kualitas dataset yang digunakan. Terdapat beberapa langkah dalam tahapan ini, seperti penghapusan kolom-kolom yang tidak digunakan dalam pemodelan, serta penanganan nilai-nilai yang hilang (*missing values*). Hasil yang didapatkan setelah proses *filter* pada tiap blok dan plot, hanya terdapat 15 baris tanaman pisang yang terjangkit penyakit *Black Sigatoka*.

#### 2. Data Transformation

*Data Transformation* merupakan proses mengubah format, struktur, ataupun nilai data ke format yang dapat dianalisis untuk pemodelan. Proses ini dilakukan secara manual untuk memastikan data terstruktur dengan baik dan sesuai dengan kebutuhan analisis serta pemodelan yang akan dilakukan.

#### 3. Data Augmentation

*Data Augmentation* merupakan proses memvariasikan dataset yang bertujuan untuk meningkatkan keberagaman data yang digunakan dalam pelatihan model, sehingga dapat meningkatkan kinerja dan akurasi model yang dibuat. *Data Augmentation* yang dilakukan yaitu mengisi kekosongan data yang terjadi pada rentang waktu tertentu, seperti salah satunya pada tanggal 1 April 2019 hingga 7 April 2019 dengan cara menambahkan data berdasarkan gejala penyakit pada hari sebelumnya.

### 3.3 Label Encoding

*Label Encoding* merupakan salah satu cara untuk mengubah variabel kategorikal menjadi representasi numerikal. *Label encoding* dilakukan karena algoritma *machine learning* dapat memproses data dan membuat model dalam bentuk data numerik saja. Proses *label encoding* dalam penelitian ini mengubah setiap nilai unik pada variabel kategorikal menjadi nilai numerik ordinal untuk memudahkan pemodelan. Kondisi daun, seperti “Belum Tumbuh, C, SR, SP, dan DF”, diubah menjadi 0, 1, 2, 3, dan 4. Proses ini menggunakan variabel *label mapping* dan fungsi *df.replace* dari *library* *pandas*. Nilai kosong dianggap sebagai kondisi “Belum Tumbuh” dan diisi dengan angka 0 menggunakan *df.fillna*. Selanjutnya, semua kolom dikonversi menjadi tipe data integer agar model dapat memprosesnya dengan mudah. Tabel 4 merupakan hasil *label encoding*.

Tabel 4 Hasil label encoding

3/31/2019	0	0	0	1	1	1	1	1	2
4/1/2019	0	0	0	1	1	1	1	1	2
4/2/2019	0	0	0	1	1	1	1	1	2
4/3/2019	0	0	0	1	1	1	1	1	2
4/4/2019	0	0	0	1	1	1	1	1	2
4/5/2019	0	0	0	1	1	1	1	1	2
4/6/2019	0	0	0	1	1	1	1	1	2
4/7/2019	0	0	0	1	1	1	1	1	2
4/8/2019	0	0	1	1	1	1	1	4	4
...	...	...	...	...	...	...	...	...	...

4/14/2019	0	0	1	1	1	2	3	4	4
...	...	...	...	...	...	...	...	...	...
7/2/2019	0	0	4	4	4	4	4	4	4

### 3.4 Split Data

Split data (pembagian data) dilakukan setelah dataset dilakukan *preprocessing* dan *label encoding*. Data dibagi menjadi dua bagian, antara lain data *training* dan data *testing*. Data *training* digunakan untuk data pelatihan model *machine learning* yang akan dibuat. Data *testing* digunakan untuk mengetahui hasil akurasi atau evaluasi dari model yang telah dibuat. Rasio split data yang dilakukan yaitu data *training* sebesar 80 % dan data *testing* sebesar 20%. Tabel 5 merupakan hasil *splitting* (pembagian) data yang dilakukan.

Tabel 5 Hasil split data

No.	Dataset	Partisi	Jumlah Data
1	Data <i>Training</i>	31 Maret 2019 - 8 Juni 2019	70 data
2	Data <i>Testing</i>	9 Juni 2019 - 2 Juli 2019	18 data

### 3.5 Pembuatan Model LSTM

Model prediksi atau peramalan dibuat menggunakan arsitektur LSTM melalui fungsi *Sequential()* dari *framework* Keras. Model ini terdiri dari *layer* yang saling terhubung secara berurutan, yang berarti *output* suatu *layer* menjadi input untuk *layer* berikutnya. Contoh arsitektur LSTM memiliki 256 *neuron hidden*. Penggunaan *neuron hidden* yang terlalu banyak dapat meningkatkan kompleksitas model, sehingga bisa menyebabkan *overfitting* saat *training*. *Training* model dalam penelitian ini dilakukan dengan melakukan perbandingan parameter *hidden neuron* dan *epoch*. Percobaan ini bertujuan untuk memperoleh model prediksi dengan hasil akurasi atau *error* yang optimal. Tabel 6 merupakan komposisi parameter yang digunakan untuk melakukan perbandingan model optimal.

Tabel 6 Komposisi training model

No.	Jumlah <i>Neuron Hidden</i>	<i>Epoch</i>
1	10	40
2		50
3		60
4	15	40
5		50
6		60
7	20	40
8		50
9		60
10	25	40
11		50
12		60
13	30	40
14		50
15		60

### 3.6 Pengujian dan Evaluasi Model

Pengujian model dilakukan setelah model dihasilkan dari proses *training*. Proses pengujian model ini akan memprediksi data *testing* dan merupakan tahapan penting dalam penggunaan model karena hasil pengujian akan memberikan gambaran tentang kinerja model dalam menggeneralisasi pada data yang belum pernah dilatih sebelumnya. Metrik evaluasi yang digunakan yaitu *Mean Absolute Error* dan *Mean Absolute Percentage Error*. Berdasarkan hasil model yang telah dibuat, model terbaik terletak pada percobaan ke-12 dengan nilai MAE sebesar 0.084 dan nilai MAPE sebesar 5.7%. Tabel 5 merupakan metrik evaluasi model.

Tabel 7 Evaluasi model

No.	Jumlah Neuron Hidden	Epoch	Metrik Evaluasi	
			MAE	MAPE
1	10	40	0.43	14.52%
2		50	0.32	13.52%
3		60	0.43	13.37%
4	15	40	1.59	14.40%
5		50	2.03	21%
6		60	1.01	17.20%
7	20	40	0.29	17.20%
8		50	0.85	37%
9		60	0.42	12.50%
10	25	40	1.16	15.14%
11		50	2.5	31%
12		60	0.084	5.70%
13	30	40	0.51	7%
14		50	0.21	8.90%
15		60	0.26	10.60%

Hasil evaluasi model menggunakan nilai MAE dan MAPE menunjukkan bahwa penggunaan *neuron hidden* sebesar 25 hingga 30 serta jumlah *epoch* 60 memberikan nilai kesalahan prediksi yang lebih kecil dibandingkan komposisi *training* lainnya. Hal ini menunjukkan bahwa komposisi tersebut memberikan keseimbangan yang optimal antara kompleksitas model dan juga durasi pelatihan. Jumlah *Epoch* 60 juga memberikan waktu yang cukup bagi model untuk menyempurnakan bobot serta bias melalui proses pelatihan yang dilakukan. Terdapat juga nilai *loss training* dan *loss validation* ketika proses *training* berlangsung. Tabel 6 merupakan nilai *loss training* yang dihasilkan.

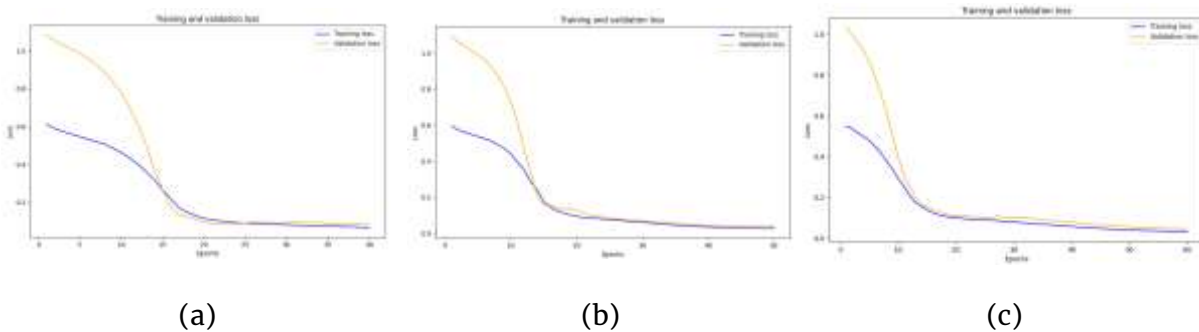
Tabel 8. Nilai *loss training* dan *loss validation*

No.	Jumlah Neuron Hidden	Epoch	Metrik Evaluasi	
			<i>Loss Training</i>	<i>Loss Validation</i>
1	10	40	0.0648	0.0897
2		50	0.0866	0.1333
3		60	0.0398	0.0562
4	15	40	0.0813	0.095
5		50	0.0727	0.0982
6		60	0.065	0.063
7	20	40	0.0628	0.0699
8		50	0.0351	0.0685
9		60	0.0484	0.0594
10	25	40	0.0667	0.0872
11		50	0.0405	0.0461
12		60	0.0234	0.0427
13	30	40	0.032	0.0485
14		50	0.0285	0.0421
15		60	0.0311	0.0442

Tabel diatas menunjukkan bahwa percobaan ke-14 memiliki nilai *loss training* dan *loss validation* masing-masing sebesar 0.0285 dan 0.0421. Nilai *training loss* menunjukkan seberapa baik model beradaptasi pada data pelatihan, sedangkan *validation loss* mencerminkan kemampuan generalisasi model pada data yang tidak terlihat selama pelatihan. *Loss* yang rendah penting karena menunjukkan model mampu mengidentifikasi pola dan generalisasi dengan baik. Meskipun percobaan

ke-14 memiliki nilai *loss* terkecil dibandingkan dengan percobaan lainnya, pemilihan model optimal tetap didasarkan pada *error* metrik (MAE dan MAPE) yang kecil, karena metrik ini lebih relevan untuk mengukur kinerja prediksi model dalam konteks nyata.

Kesimpulan yang dapat diambil dari hasil analisis bahwa nilai *loss* rendah tidak selalu berarti performa model optimal. Model yang terlalu kompleks dapat mengurangi *loss* secara signifikan, tetapi gagal dalam generalisasi, sehingga model yang lebih sederhana seringkali lebih baik untuk prediksi data baru meskipun *loss*-nya sedikit lebih besar. Berikut ini merupakan gambaran terkait grafik *loss* percobaan ke-13, 14, dan 15 dengan *neuron hidden* sebesar 25 dan *epoch* masing-masing sebesar 40, 50, dan 60.



Gambar 2 Grafik nilai *training* dan *validation loss*; (a) Percobaan ke-13, (b) Percobaan ke-14, (c) Percobaan ke-15.

Grafik *training* dan *validation loss* menunjukkan bahwa dengan jumlah *epoch* 40 dan 60 nilai *loss* yang dihasilkan cenderung naik pada awal atau akhir pelatihan, dikarenakan model belum berhasil menemukan pola-pola yang signifikan pada *dataset*. Namun nilai *loss* mengalami penurunan sedikit demi sedikit seiring bertambahnya *epoch*. Akan tetapi, *epoch* 40 dan 60 masih menghasilkan nilai *loss* yang fluktuasi yang menunjukkan model belum memahami data sepenuhnya. Grafik *loss* pada jumlah *epoch* 50 menghasilkan nilai *loss* yang menurun secara stabil, mulai dari awal hingga akhir pelatihan. Berdasarkan perbandingan tersebut dapat disimpulkan bahwa dengan meningkatkan jumlah data dapat menghasilkan nilai *error* dan nilai *loss* yang kecil, sehingga model akan mempelajari pola data dengan baik, memahami kondisi yang terjadi, serta dapat melakukan prediksi di lapangan secara nyata dan benar.

#### 4. KESIMPULAN

Hasil dari metode LSTM dalam melakukan prediksi penyakit *Black Sigatoka*, menghasilkan model dengan nilai *error* yang kecil. *Hyperparameter* berupa *neuron hidden* dan *epoch* yang disesuaikan pada tiap percobaan berpengaruh terhadap kinerja model, dengan menghasilkan nilai *error* terkecil pada percobaan ke-12. Hasil evaluasi didapatkan nilai MAE dan MAPE masing-masing sebesar 0.084 dan 5.7% yang menunjukkan bahwa model dapat membaca pola dan memiliki nilai *error* yang kecil.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih setinggi-tingginya kepada Institut Teknologi Sumatera atas bantuan hibah penelitian (No. 1539az/IT9.2.1/PT.01.03/2024) melalui program “Hibah Penelitian Itera 2024”.

#### DAFTAR PUSTAKA

- [1] Kalsum, U., Subandi, Y., & Wiratma, H. D. (2023). PETANI TANGGAMUS MITRA PT. GREAT GIANT PINEAPPLE MENGEKSPOR PISANG MAS KE SINGAPURA TAHUN 2021. PRIMER : Jurnal Ilmiah Multidisiplin, 1(2), 152–164. <https://doi.org/10.55681/primer.v1i2.63>
- [2] Badan Pusat Statistik. (2022). Statistik Hortikultura 2022 ( dan P. Direktorat Statistik Tanaman Pangan, Hortikultura (ed.)). BPS RI/BPS-Statistics Indonesia.



- [3] Bebbber, D. P. (2019). Climate change effects on Black Sigatoka disease of banana. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1775), 20180269. <https://doi.org/10.1098/rstb.2018.0269>
- [4] Yoka, M. F., & Wijayanto, S. (2021). Forecasting Penjualan Gas LPG di Toko Sembako Menggunakan Metode Fuzzy Time Series. *JUPITER: Jurnal Penelitian Ilmu Dan Teknologi Komputer*, 13(2), 87–96. <https://doi.org/10.5281/3541.jupiter.2021.10>
- [5] Gala, I. G., & Kulkarni, P. (2023). A comprehensive review of classical and deep learning-based time series models. In *AI-Based Metaheuristics for Information Security and Digital Media*. <https://doi.org/10.1201/9781003107767-7>
- [6] Lasijan, T. G., Santoso, R., & Hakim, A. R. (2023). PREDIKSI HARGA EMAS DUNIA MENGGUNAKAN METODE LONG-SHORT TERM MEMORY. *Jurnal Gaussian*, 12(2), 287–295. <https://doi.org/10.14710/j.gauss.12.2.287-295>
- [7] Aisyah, D., Purboyo, T. W., & Kallista, M. (2023). Prediksi Penderita Tuberkulosis Dengan Algoritma Long Short-Term Memory. *E-Proceeding of Engineering*, 10(1).
- [8] Ningrum, A. A., Syarif, I., Gunawan, A. I., Satriyanto, E., & Muchtar, R. (2021). Algoritma Deep Learning-LSTM untuk Memprediksi Umur Transformator. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(3). <https://doi.org/10.25126/jtiik.2021834587>
- [9] Aprian, B. A., Azhar, Y., & Nastiti, V. R. S. (2020). Prediksi Pendapatan Kargo Menggunakan Arsitektur Long Short Term Memory. *Jurnal Komputer Terapan*, 6(2). <https://doi.org/10.35143/jkt.v6i2.3621>