

# PREDIKSI CURAH HUJAN DI KABUPATEN BADUNG, BALI MENGGUNAKAN METODE LONG SHORT-TERM MEMORY

1) Brando Dharma Saputra , 2) Lely Hiryanto , 3) Teny Handhayani

1) Teknik Informatika, FTI, Universitas Tarumanaraga  
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia  
email : [brando.535180079@stu.untar.ac.id](mailto:brando.535180079@stu.untar.ac.id)

2) Teknik Informatika, FTI, Universitas Tarumanaraga  
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia  
email : [lelyh@fti.untar.ac.id](mailto:lelyh@fti.untar.ac.id)

3) Teknik Informatika, FTI, Universitas Tarumanaraga  
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia  
email : [tenyh@fti.untar.ac.id](mailto:tenyh@fti.untar.ac.id)

## ABSTRACT

*Rainfall is the height of rainwater that falls on a flat area, assuming it doesn't evaporate, doesn't seep, and doesn't flow. Rain levels are measured in mm (millimeters). The target of the research being conducted is in Badung Regency, Bali because Bali is a tourist area that is often visited by tourists and from Indonesian itself, so predictions of meteorology, such as rainfall will greatly impact tourism. In this test, predictions use the Long Short Term Memory (LSTM) method, using daily weather data from the BMKG from 2010 to 2020 as training data and daily weather data for 2021 as prediction data. Based on the test results above, the results show that the two LSTM tests with LSTM Model 128.64 and LSTM Model 64.32 have low MAE and MAPE error values. From First Scenario, the Mean Absolute Error (MAE) value is 8.97246598930908 and Mean Absolute Percentage Error (MAPE) value is 1.7657206683278308%. From Second Scenario, the Mean Absolute Error is 9.706669940783014 and Mean Absolute Percentage Error is 1.9028466692362323%. From the MAE and MAPE values obtained in these two scenarios, it can be proven that from the evaluation results of Rainfall predictions in Badung Regency, Bali, the predictions can be said to be very accurate because they have an error value of less than 10.*

## Key words

*Badung Regency , Bali , Meteorology , Long Short Term Memory , Rainfall Prediction*

## 1. Pendahuluan

Curah hujan adalah ketinggian air hujan yang jatuh pada tempat yang datar dengan asumsi tidak menguap, tidak meresap dan tidak mengalir. Tingkat hujan yang diukur dalam satuan mm (milimeter) adalah air hujan setinggi 1 mm yang jatuh pada tempat yang datar dengan

luas 1 m<sup>2</sup> dengan asumsi tidak ada yang menguap, mengalir dan meresap [1]. Data curah hujan penting untuk perencanaan teknik, terutama untuk sistem drainase seperti irigasi, bendungan, drainase perkotaan, pelabuhan, dermaga, dan struktur air lainnya.

Long Short-Term Memory (LSTM) merupakan salah satu bentuk RNN yang paling umum yang dimaksudkan untuk menghindari masalah ketergantungan jangka panjang dan sesuai untuk memproses serta memprediksi deret waktu [2]. LSTM dapat digunakan untuk prediksi, karena dapat memproses dan memprediksi data yang sudah disimpan dalam jangka waktu yang lama. Untuk memprediksi Curah Hujan, terdapat beberapa faktor yang dapat mempengaruhi, seperti Temperatur, Kecepatan Angin, Kelembapan, dan lain-lain [3]. Untuk itu, digunakan metode LSTM karena dapat memproses perhitungan prediksi dengan banyak variabel dan dapat menghindari masalah ketergantungan jangka panjang dan sesuai untuk memproses deretan waktu yang banyak.

Dilakukan proses prediksi Curah Hujan menggunakan metode Long Short Term Memory di Bali karena Bali merupakan pusat wisata di Indonesia yang paling terkenal di semua kalangan , baik WNI maupun orang dari negara asing (turis) . Diharapkan dengan adanya penelitian untuk memprediksi Curah Hujan di Kabupaten Badung , Bali dapat membantu orang untuk membuat jadwal wisata lebih baik sehingga wisata berjalan dengan lancar.

## 2. Metode Penelitian

Fokus dari penelitian ini adalah untuk membuat prediksi tentang metereologi, yaitu kadar curah hujan di Kabupaten Badung, Bali menggunakan metode Long Short-Term Memory. Proses penelitian dilakukan dengan pengumpulan data harian BMKG dari tahun

2010 sampai tahun 2021. Kemudian dilakukan Pre-Processing data dengan guna menghilangkan missing value dari data. Kemudian dilakukan proses prediksi menggunakan metode LSTM. Dan akhirnya, dilakukan validasi menggunakan Evaluation Metrics guna menghitung nilai error dari hasil penelitian.

2.1. Meteorologi

Meteorologi adalah ilmu interdisipliner yang mempelajari masalah atmosfer, misalnya, suhu, udara, cuaca, angin, dan berbagai sifat fisika dan kimia atmosfer lainnya yang digunakan untuk keperluan prakiraan cuaca. Dalam kamus besar bahasa Indonesia, meteorologi di definisikan sebagai cabang ilmu geografi yang mempelajari tentang ciri-ciri fisik dan kimia atmosfer untuk meramalkan keadaan cuaca di suatu tempat secara khusus dan di seluruh dunia secara umum [4].

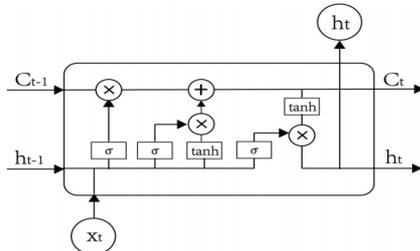
2.2. Pre-Processing Data

Pre-Processing Data adalah teknik awal data mining untuk mengubah data mentah menjadi format dan informasi yang lebih efisien dan bermanfaat. Format pada data mentah yang diambil dari berbagai macam sumber sering sekali mengalami error, missing value, dan tidak konsisten. Sehingga, perlu dilakukan pembenahan format agar hasil menjadi akurat [5].

2.3. Long Short Term Memory (LSTM)

Model LSTM menyaring informasi melalui struktur gerbang untuk mempertahankan dan memperbarui keadaan sel memori. Struktur pintunya mencakup input gate, forget gate, dan output gate. Setiap sel memori memiliki tiga lapisan sigmoid dan satu lapisan tanh. Gambar 1 menunjukkan struktur metode LSTM [6].

Lapisan sigmoid menampilkan angka antara nol dan satu, menggambarkan berapa banyak masing-masing komponen yang harus dilewati. Nilai nol berarti “jangan biarkan apa pun lewat” sementara nilai satu berarti “biarkan semuanya lewat”. Sigmoid akan bernilai nol jika terletak pada rentang  $0 \leq \sigma \leq 0,5$  dan bernilai satu jika terletak pada rentang  $0,5 < \sigma \leq 1$ .



Gambar 1 Struktur Model LSTM

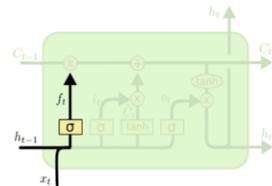
Fungsi sigmoid dan tanh adalah sebagai berikut :

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

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

Keterangan :

- $\sigma$  = sigmoid
- $x$  = nilai input
- $e$  = konstanta matematika



$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Gambar 2 Model Forget Gate LSTM

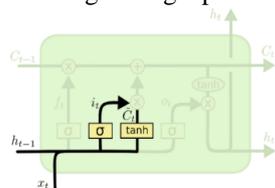
Forget gate dalam unit LSTM menentukan informasi status sel mana yang dibuang dari model. Seperti yang ditunjukkan di gambar 2, sel memori menerima output  $h_{t-1}$  dari momen sebelumnya dan informasi  $x_t$  dari momen saat ini sebagai input dan menggabungkannya dalam vektor panjang  $[h_{t-1}, x_t]$ . Notasi  $[h_{t-1}, x_t]$  merupakan operasi konkatenasi, artinya menambahkan baris dari  $x_t$  dengan baris-baris dari  $h_{t-1}$  melalui transformasi  $\sigma$  menjadi [7] :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad \dots(3)$$

Keterangan :

- $f_t$  = Forget gate
- $\sigma$  = fungsi sigmoid
- $W_f$  = nilai weight untuk Forget gate
- $h_{t-1}$  = nilai output sebelum orde ke-t
- $x_t$  = nilai input pada orde ke-t
- $b_f$  = nilai bias pada Forget gate

Fungsi utama dari Forget gate adalah untuk merekam berapa banyak status sel  $C_{t-1}$  dari waktu sebelumnya dicadangkan ke status sel  $C_t$  dari waktu saat ini. Gerbang akan menampilkan nilai anatara 0 dan 1 berdasarkan  $h_{t-1}$  dan  $x_t$ , dimana 1 menunjukkan pembuangan lengkap.



$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

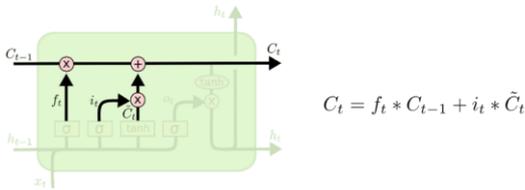
Gambar 3 Model Input Gate LSTM

Gerbang input akan menunjukkan berapa banyak input jaringan waktu saat ini,  $x_t$  dicadangkan ke dalam status sel  $C_t$ , yang mencegah konten tidak signifikan dari memasuki sel memori. Input gate memiliki 2 fungsi, yang pertama adalah menemukan keadaan sel yang harus diperbarui; nilai yang akan diperbarui dipilih oleh lapisan sigmoid, seperti persamaan di bawah ini [8] :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \dots(4)$$

Keterangan :

- $i_t$  = Input gate
- $\sigma$  = fungsi sigmoid
- $W_i$  = nilai weight untuk Input gate
- $h_{t-1}$  = nilai output sebelum orde ke-t
- $x_t$  = nilai input pada orde ke-t
- $b_i$  = nilai bias pada Input gate



Gambar 4 Model Cell State LSTM

Sedangkan fungsi lainnya adalah untuk memperbarui informasi ke b diperbau ke keadaan sel. Vektor kandidat baru  $C_t$  dibuat melalui tanh untuk mengontrol berapa banyak informasi baru yang ditambahkan, seperti dalam persamaan dibawah ini, digunakan untuk memperbarui keadaan dari sel ke sel memori [9].

$$c_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad \dots(5)$$

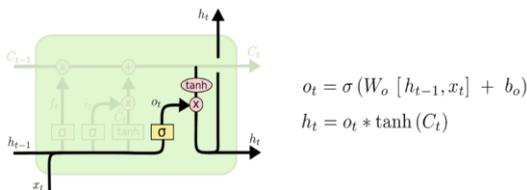
Keterangan :

- $c_t$  = nilai baru yang ditambahkan ke Cell state
- tanh = fungsi tanh
- $W_c$  = nilai weight untuk Cell state
- $h_{t-1}$  = nilai output sebelum orde ke-t
- $x_t$  = nilai input pada orde ke-t
- $b_c$  = nilai bias pada Cell state

$$C_t = (f_t \cdot C_{t-1}) + (i_t \cdot c_t) \quad \dots(6)$$

Keterangan :

- $C_t$  = Cell state
- $f_t$  = Forget gate
- $C_{t-1}$  = Cell state sebelum orde ke-t
- $i_t$  = Input gate
- $c_t$  = nilai baru yang ditambahkan ke Cell state



Gambar 5 Model Output Gate LSTM

Gerbang output mengontrol berapa banyak keadaan sel saat ini dibuang. Informasi keluaran pertama-tama ditentukan oleh lapisan sigmoid, kemudian keadaan sel diproses oleh tanh dan dikalikan dengan keluaran lapisan sigmoid untuk mendapatkan bagian di keluaran akhir [10].

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \dots(7)$$

Keterangan :

- $O_t$  = Output gate
- $\sigma$  = fungsi sigmoid
- $W_o$  = nilai weight untuk Output gate
- $h_{t-1}$  = nilai output sebelum orde ke-t
- $x_t$  = nilai input pada orde ke-t
- $b_o$  = nilai bias pada Output gate

$$h_t = O_t \cdot \tanh(C_t) \quad \dots(8)$$

Keterangan :

- $h_t$  = nilai output orde ke-t
- $O_t$  = Output gate
- tanh = fungsi tanh
- $C_t$  = Cell state

## 2.4. Evaluation Metrics

Metrik evaluasi yang digunakan pada penelitian ini adalah MAE (Mean Absolute Error) dan MAPE (Mean Absolute Percentage Error). MAE menghitung rata-rata absolut dari kesalahan prediksi yang tidak terlalu terpengaruhi oleh data outlier [11]. Sedangkan MAPE merupakan nilai rata-rata persentase kesalahan atau error dari beberapa periode [12]. Perhitungan MAE dan MAPE dapat dilihat dalam rumus berikut ini:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - Y'_i| \quad \dots(9)$$

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \frac{Y_i - Y'_i}{Y_i} \cdot 100 \% \quad \dots(10)$$

Keterangan :

- MAE = Mean Absolute Error
- MAPE = Mean Absolute Percentage Error
- n = banyak data
- $Y_i$  = Nilai data aktual di periode ke-i
- $Y'_i$  = Nilai data prediksi di periode ke-i

Berikut klasifikasi nilai MAE dilihat dalam Tabel 1 :

Tabel 1 Klasifikasi Nilai MAE [11]

No	Nilai MAE	Pengertian
1	$0 < MAE < 10$	Sangat Akurat
2	$10 < MAE < 20$	Baik
3	$20 < MAE < 50$	Cukup
4	$MAE > 50$	Tidak Akurat

Berikut klasifikasi nilai MAPE dilihat dalam Tabel 2 :

Tabel 2 Klasifikasi Nilai MAPE [12]

No	Nilai MAPE	Pengertian
1	$0 < MAPE < 10$	Sangat Akurat
2	$10 < MAPE < 20$	Baik
3	$20 < MAPE < 50$	Cukup
4	$MAPE > 50$	Tidak Akurat

### 3. Hasil Percobaan

Berikut model LSTM yang akan digunakan untuk pengujian :

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 14, 128)	70144
lstm_1 (LSTM)	(None, 64)	49408
dropout (Dropout)	(None, 64)	0
dense (Dense)	(None, 1)	65

```

=====
Total params: 119,617
Trainable params: 119,617
Non-trainable params: 0
    
```

Gambar 6 Model LSTM 128 , 64 (Skenario 1)

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 14, 64)	18688
lstm_1 (LSTM)	(None, 32)	12416
dropout (Dropout)	(None, 32)	0
dense (Dense)	(None, 1)	33

```

=====
Total params: 31,137
Trainable params: 31,137
Non-trainable params: 0
    
```

Gambar 7 Model LSTM 64 , 32 (Skenario 2)

Kedua model ini menggunakan hyperparameter yang sama yaitu optimizer adam, activation function tanh, epoch sebesar 50, batch size sebesar 16, validation split sebesar 0.2 dan verbose sebesar 1.

Berikut hasil pengujian skenario 1 dan 2 :

Skenario 1 :

- LSTM Model : 128 , 64
- MAE : 8.97246598930908
- MAPE : 1.7657206683278308 %

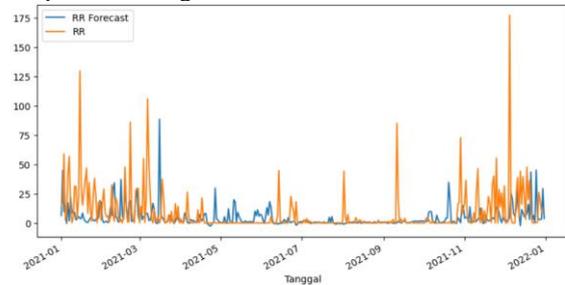
Skenario 2 :

- LSTM Model : 64 , 32
- MAE : 9.706669940783014
- MAPE : 1.9028466692362323 %

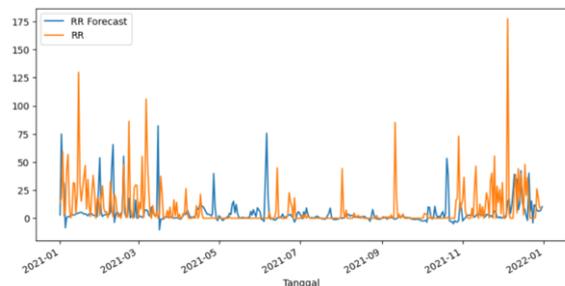
Berdasarkan hasil pengujian diatas, didapatkan hasil bahwa kedua pengujian LSTM dengan LSTM Model 128,64 dan LSTM Model 64,32 memiliki nilai MAE dan MAPE yang rendah. Dari Skenario 1 didapat MAE sebesar 8.97246598930908 dan MAPE sebesar 1.7657206683278308 %. Dari Skenario 2 didapat MAE sebesar 9.706669940783014 dan MAPE sebesar 1.9028466692362323 %. Dari nilai MAE dan MAPE yang didapat pada kedua scenario tersebut, dapat

dibuktikan bahwa dari hasil evaluasi prediksi Curah Hujan di Kabupaten Badung, Bali, prediksi bisa dikatakan sangat akurat karena memiliki nilai error kurang dari 10.

Berikut grafik hasil prediksi Curah Hujan di Kabupaten Badung, Bali :



Gambar 8 Grafik Hasil Prediksi (Skenario 1)



Gambar 9 Grafik Hasil Prediksi (Skenario 2)

Berdasarkan grafik prediksi Curah Hujan di Kabupaten Badung, Bali di Gambar 6, hasil prediksi tidak begitu menyimpang dengan data actual. Prediksi dilakukan di tahun 2021 dan dibandingkan dengan data aktual. Di dalam grafik, grafik prediksi digambarkan dengan warna biru dan grafik data actual digambarkan dengan warna jingga. Dari grafik, dapat disimpulkan bahwa metode Long Short-Term Memory adalah metode yang baik untuk memprediksi dengan jangka waktu yang jauh.

### 4. Kesimpulan

Berdasarkan hasil pengujian diatas, didapatkan hasil bahwa kedua pengujian LSTM dengan LSTM Model 128,64 dan LSTM Model 64,32 memiliki nilai MAE dan MAPE yang rendah. Dari Skenario 1 didapat MAE sebesar 8.97246598930908 dan MAPE sebesar 1.7657206683278308 %. Dari Skenario 2 didapat MAE sebesar 9.706669940783014 dan MAPE sebesar 1.9028466692362323 %. Dari nilai MAE dan MAPE yang didapat pada kedua scenario tersebut, dapat dibuktikan bahwa dari hasil evaluasi prediksi Curah Hujan di Kabupaten Badung, Bali, prediksi bisa dikatakan sangat akurat karena memiliki nilai error kurang dari 10.

Berdasarkan grafik prediksi Curah Hujan di Kabupaten Badung, Bali, hasil prediksi tidak begitu menyimpang dengan data actual. Prediksi dilakukan di tahun 2021 dan dibandingkan dengan data actual. Di dalam grafik, grafik prediksi digambarkan dengan warna biru dan grafik data actual digambarkan dengan warna jingga. Dari grafik, dapat disimpulkan bahwa metode Long Short-Term Memory adalah metode yang baik untuk memprediksi dengan jangka waktu yang jauh.

## REFERENSI

- [1] Zhang, Chang-Jiang ; Zeng, Jing ; Wang, Hui-Yuan ; Ma, Lei-Ming ; Chu, Hai . “Correction model for rainfall forecasts using the LSTM with multiple meteorological factors” . Meteorological Application . Vol.27 , No.1 . November 2019.
- [2] Wiranda, Laras ; Sadikin, Mujino . “Penerapan Long Short-Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk PT. Metiska Farma” . Jurnal Nasional Pendidikan Teknik Informatika . Vol. 8 , No. 3 . 2019.
- [3] Haq, Dina Zatusiva ; Novitasari, Dian Candra Rini ; Hamid, Abdulloh ; Ulinnuha, Nurissaidah, Arnita ; Farida, Yuniar ; Nugraheni, RR. Diah ; Nariswari, Rinda, Ilham, Rohayani, Hetty ; Pramulya, Rahmat ; Widjayanto, Ari . “Long Short-Term Memory Algorithm for Rainfall Prediction Based on El -Nino and IOD Data”. Procedia Computer Science . Vol. 179 . 2021.
- [4] Owen, Michael ; Vincent, Vincent ; Ambarita, Riama Br ; Indra, Evta . “Implementasi Metode Long Short-Term Memory Untuk Memprediksi Pergerakan Nilai Harga Emas” , Jurnal TEKNINKOM . Vol. 5 , No. 1 . Juni 2022.
- [5] Lindemann, Benjamin ; Muller, Timo ; Vietz, Hannes ; Jazdi, Nasser ; Weyrich, Michael . “A survey on long short-term memory networks for time series prediction” . Procedia CIRP 99 . Vol. 99 . 2021.
- [6] Ouma, Yashon O. ; Cheruyot, Rodrick ; Wachera, Alice N. . “Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: case study of Nzoia hydrologic basin” . Complex & Intelligent Systems . Vol 8 . 2022.
- [7] Tang, Qi ; Shi, Ruchen ; Fan, Tongmei ; Ma, Yidan ; Huang, Jingyan . “Prediction of Financial Time Series Based on LSTM Using Wavelet Transform and Singular Spectrum Analysis” . Hindawi (Mathematical Problems in Engineering) . Vol. 2021 . Juni 2021.
- [8] Tao, Qing ; Liu, Fang ; Li, Yong ; Sidorov, Dennis . “Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU” . IEEE Access . Vol. 7 . 2019.
- [9] Ulumuddin, Ihyak ; Sunardi ; Fadlil, Abdul . “BITCOIN PRICE PREDICTION USING LONG SHORT-TERM MEMORY (LSTM)” . Jurnal Mantik . Vol. 4 , No. 2 . Agustus 2020.
- [10] Rizki, Muhammad ; Basuki, Setio ; Azhar, Yufis . “Implementasi Deep Learning Menggunakan Arsitektur Long Short-Term Memory Untuk Prediksi Curah Hujan Kota Malang” . REPOSITOR . Vol. 2 , No. 3 . 2020.
- [11] Badriyah, Jamilatul ; Fariza, Arna ; Harsono, Tri . “Prediksi Curah Hujan Menggunakan Long Short Term Memory” . Jurnal Media Informatika Budidarma . Vol. 6 No. 3 . Juli 2022.
- [12] Rayendra, Tezhar Trastaronny Pastika Nugraha . “Evaluasi Prediksi Curah Hujan dengan Algoritma Backpropagation di BMKG Cilacap” . ZONAsi Jurnal Sistem Informasi . Vol. 2 , No. 2 . September 2020.

**Brando Dharma Saputra**, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.

**Lely Hiryanto**, memperoleh gelar S.T. dari Universitas Tarumanagara pada tahun 2001. Kemudian memperoleh gelar M.Sc. dari Curtin University of Technology pada tahun 2006. Kemudian memperoleh gelar Ph.D. dalam bidang Ilmu Komputer dari Curtin University of Technology pada tahun 2022. Saat ini aktif sebagai Dosen Tetap di Fakultas Teknologi Informasi, Universitas Tarumanagara, Jakarta.

**Teny Handhayani**, memperoleh gelar S.Kom. dari Institut Pertanian Bogor tahun 2008. Kemudian memperoleh gelar Magister Ilmu Komputer dari Universitas Indonesia tahun 2013, dan kemudian memperoleh gelar Philosophy, Ilmu Komputer dari University of York tahun 2021. Saat ini aktif sebagai Dosen Tetap di Fakultas Teknologi Informasi, Universitas Tarumanagara, Jakarta.