

Multivariate Forecasting Curah Hujan Menggunakan Algoritma LSTM Di Kota Denpasar

Putu Sugiartawan¹, Servus Gusprio Santoso²

Fakultas Teknologi dan Informasi
Institut Bisnis dan Teknologi Indonesia
Denpasar, Bali

e-mail: ¹putu.sugiartawan@instiki.ac.id, ²servus.santoso@instiki.ac.id

Abstrak

Badan Meteorologi Klimatologi dan Geofisika (BMKG) adalah institusi pemerintah yang bertugas untuk mencatat informasi curah hujan. Fluktuasi data curah hujan di Indonesia berdampak pada kebutuhan akan data curah hujan yang akurat kedepannya. Curah hujan merupakan suatu iklim yang sangat berpengaruh dalam mencirikan kondisi iklim di Indonesia, misalnya bidang pertanian, industri, dan transportasi. Untuk mengetahui curah hujan di Indonesia memiliki keragaman khususnya di Kota Denpasar, diperlukan adanya Peramalan atau forecasting yang dapat mengetahui pola tingkat keakuratan curah hujan perbulan di kantor BMKG. Teknik forecasting bertujuan untuk memprediksi pola data yang sebelumnya, sehingga dapat diketahui pola data selanjutnya. Pada penelitian ini menggunakan teknik recurrent neural network (RRN) yang digunakan dalam penelitian ini yaitu: Long Short Term Memory (LSTM). Dalam memprediksi tingkat Keakuratan curah hujan perbulanya. Dengan adanya penelitian ini data yang digunakan untuk melakukan peramalan yaitu: data curah hujan perbulan dalam waktu lima tahun terakhir yaitu dari tahun 2016-2021 dari kantor BMKG kota Denpasar. Dan hasil yang diperoleh dalam melakukan peramalan ini yang menggunakan model LSTM adalah MAE data testing sebesar 166.56504 MAPE data test sebesar 11.89371%, dan MAE data training sebesar 199.19741, MAPE data training sebesar 2.518% nilai yang diukur merupakan nilai eror dengan teknik MAE dan MAPE. Arsitektur LSTM yang di gunakan terdiri dari jumlah unit neuron pada hidden layer sebanyak 30 unit, menggunakan windows size sebanyak 6.

Kata kunci: Multivariate forecasting, curah hujan, Algoritma LSTM

Abstract

The Meteorology, Climatology, and Geophysics Agency (BMKG) is the official agency that provides and distributes rainfall information. The need for the provision and service of rainfall data is increasing, demanding BMKG to be able to prepare and deliver data quickly, accurately, and precisely. Because rainfall in Indonesia has a variety of weather conditions that significantly affect life, therefore, with this research, forecasting techniques can find out the pattern of the accuracy of monthly rainfall at the BMKG office. Forecasting techniques aim to predict previous data patterns so that predictions can know different data patterns. This study's recurrent neural network (RRN) technique is Long Short Term Memory (LSTM). In predicting the level of accuracy of monthly rainfall. With this research, the data used for forecasting are monthly rainfall data in the last five years, namely from 2016-2021, from the BMKG office in Denpasar. And the results obtained in doing this forecasting using the LSTM model are MAE data testing of 166.56504 MAPE test data of 11.89371%, MAE of training data of 199.19741, MAPE data training of 2.518%. The measured value is an error value using the MAE technique and MAPE. The LSTM architecture used consists of 30 units of neurons in the hidden layer, using a window size of 6.

Keywords: Multivariate forecasting, rainfall, LSTM Algorithm

1. Pendahuluan

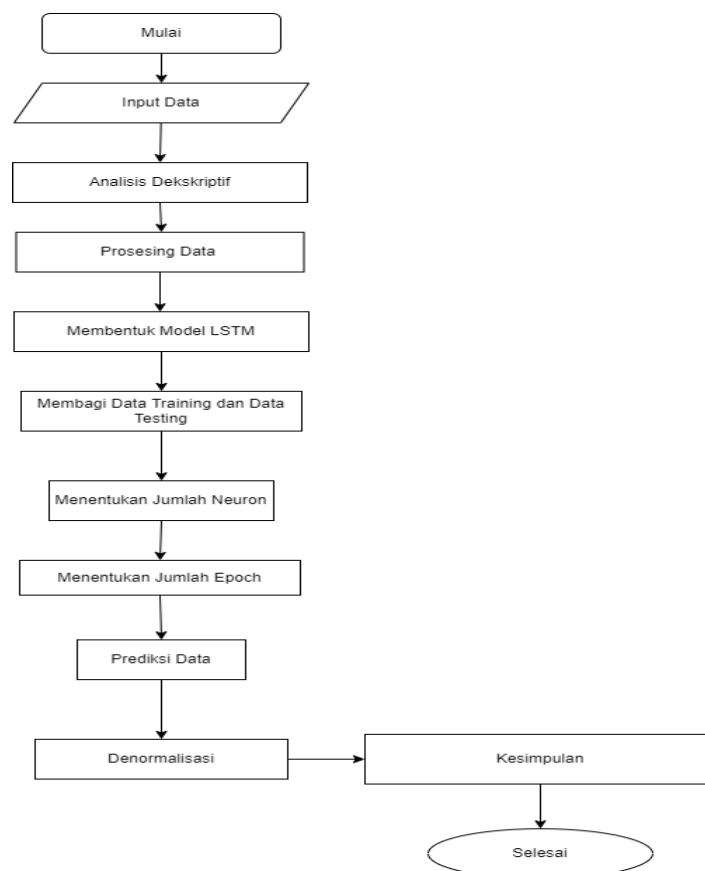
Pemanasan global berdampak pada curah hujan yang ada di Indonesia khususnya Kota Denpasar, kebutuhan akan informasi curah hujan menjadi sebuah kebutuhan di masyarakat kota Denpasar [1]. Data curah hujan banyak masyarakat yang membutuhkan dan kebutuhan tersebut semakin meningkat [2]. Beberapa faktor yang sering terjadi bahwa keadaan dari iklim, cuaca, curah hujan, angin, yang terjadi di Indonesia tidak selamanya akurat atau sesuai dengan pola setiap pada musimnya [3]. Seperti curah hujan

kadang intensitas curah hujan berbeda-beda setiap tahunnya. Begitupun yang terjadi pada kota Denpasar , dimana intensitas curah hujanya cukup rendah karena dalam satu pekan pasti terjadi hujan dengan intensitas berbeda-beda setiap bulanya. Teknik peramalan merupakan salah satu penyelesaian masalah dalam upaya penyediaan data cuaca [4], [5].

Metode yang sesuai dibutuhkan untuk melakukan teknik prediksi curah hujan. Metode yang digunakan untuk teknik prediksi adalah dengan menggunakan metode LSTM (LongShort Term Memory), metode tersebut menghasilkan nilai akurasi yang lebih tinggi dibandingkan metode lainnya [6]. Forget gate merupakan kelebihan yang dimiliki oleh LSTM [7], [8]. Penelitian akan menitik beratkan pada prediksi curah hujan untuk setiap bulanan pada Kota Denpasar menggunakan metode LSTM, dimana di kantor BMKG Stasiun Geofisika Sanglah Denpasar ini masih mengalami terjadinya anomali cuaca atau keganjilan atau keanehan yang menyimpang dari keadaan biasanya. Sehingga dari pihak BMKG tidak memastikan cuaca hujan yang ada dikota Denpasar karena adanya kendala dalam perkiraan cuaca yang masih kurang akurat. Oleh karena itu dengan adanya penelitian ini Teknik peramalan atau forecasting dapat mengetahui pola tingkat keakuratan curah hujan perbulan di kantor BMKG. Teknik forecasting bertujuan untuk memprediksi pola data yang sebelumnya, sehingga dapat diketahui pola data selanjutnya. Pada penelitian ini menggunakan teknik recurrent neural network dalam memprediksi tingkat Keakuratan curah hujan perbulanya. Salah satu teknik recurrent neural network (RNN) [9] yang digunakan dalam penelitian ini adalah Long Short Term Memory (LSTM). Model ini lebih baik dibandingkan dengan model RNN sederhana. Dan data yang digunakan dalam penelitian ini adalah: data curah hujan perbulan, selama lima tahun terakhir yaitu dari tahun 2016-2021. Data curah hujan dalam bentuk decimal, pada tahun 2018 data curah hujan tertinggi ditunjukkan dengan nilai 515.6 pada bulan Januari. Kode data curah hujan disingkat dengan kode RRR.

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah *Long Short Term Memory (LSTM)* dan penggunaan data untuk proses prediksi curah hujan, dimana data diperoleh pada kantor BMKG Stasiun Geofisika Sanglah Denpasar. Data yang diambil merupakan data curah hujan kota Denpasar dari tahun 2016 sampai 2021.



Gambar 1. Flowchart prediksi curah hujan

Penelitian yang dituangkan dalam diagram flowchart pada Gambar 1. Menggambarkan proses penelitian yang akan ditempuh sekaligus menggambarkan penelitian secara keseluruhan. Dibawah ini merupakan uraian langkah dari penelitian ini

1. Mulai. Tahap ini merupakan tahapan subjek (populasi) di Kantor BMKG Stasiun Geofisika Sanglah Denpasar. Objek yang di ambil yaitu Multivariate Ferecasting Curah Hujan Dengan Menggunakan Algoritma LSTM di Kota Denpasar.
2. Input data curah hujan yang diperoleh dari BMKG dengan kode curah hujan RRR
3. Preprocessing Data, Normalisasi data dengan metode min-max, dengan interval data diantara 0-1.
4. Analisis Deskriptif data bulanan Ethereum, Melakukan analisis Multivariat LSTM dengan tahapan berikut ini :
 - a. Membagi data menjadi data training dan testing
 - b. Menentukan jumlah neuron pada hidden layer
 - c. Menentukam jumlah epoch
 - d. Melakukan Prediksi
5. Denormalisasi Data
6. Menghitung nilai error dengan MSE
7. Kesimpulan , membuat kesimpulan dari hasil penelitian yang menggunakan algoritma LSTM, Gambar 1 merupakan flowchart dari langkah penelitian.

2.1 Normalisasi dan Denormalisasi

Tujuan dari normalisasi data adalah untuk proses penskalan nilai atribut dari data sehingga berada pada range tertentu. Salah satu metode normalisasi yang dapat di gunakan adalah Min-max normalisasi. Proses normalisasi tersebut merupakan metode dengan melakukan transformasi linear terhadap data aslinya [10], [11]. Adapun persamaan yang dapat di digunakan untuk menghitung nilai normalisasi seperti yang ditunjukkan pada persamaan satu di bawah ini.

$$X' = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Keterangan:

X' = Data hasil normalisasi, x_i = Data ke i

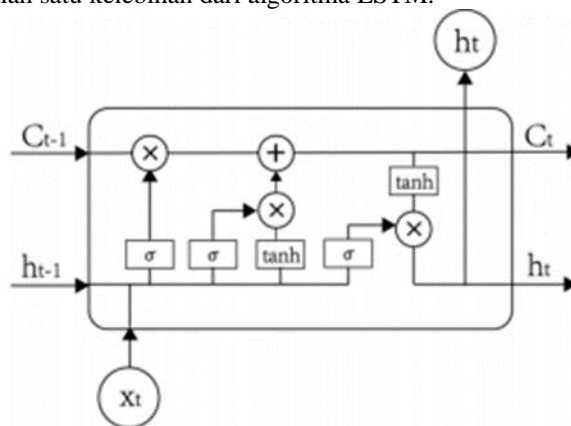
x_{min} = Data dengan nilai minimum, x_{max} = Data dengan nilai maksimum

Proses dernormalisasi data yang bertujuan untuk mengembalikan hasil keluaran jaringan agar berada pada *range* sebelumnya. Denormalisasi data dapat dilakukan dengan menggunakan persamaan di bawah ini[12]

$$x_i = X'(x_{max} - x_{min}) + x_{min}$$

2.2 Long Short Trem Memory (LSTM)

Kemampuan mengingat dari sekuens long term (ukuran data) yang sulit dicapai dengan teknik fitur tradisional merupakan salah satu kelebihan dari algoritma LSTM.



Gambar 2. Struktur LSTM

LSTM dapat menggunakan ukuran data yang lebih besar dan menggunakan semua informasi data sebagai masukan. Arsitektur LSTM di awali dari Layer pertama dengan jumlah unit dan diikuti layer LSTM lain dengan jumlah unit. Sedangkan pada lapisan akhir adalah layer keluaran dengan aktivasi single unit dan sigmoid.

3. Hasil dan Pembahasan

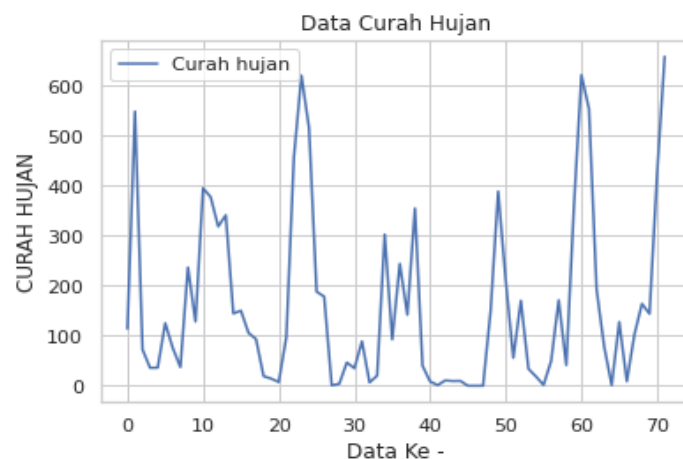
3.1. Processing Data

Pada tahap ini dilakukan pengolahan data yang nantinya akan digunakan dalam melakukan peramalan. Dan data yang diolah untuk melakukan peramalan ini merupakan data curah hujan perbulan lima tahun terakhir dari kantor BMKG Denpasar, yaitu data dari bulan Januari tahun 2016 sampai dengan bulan Desember 2021. Dan pada penelitian ini data training sebesar 70% dari total data dan data testingnya sebanyak 30% dari total data. Pembagian jumlah data training dan testing berdasarkan pada penelitian [13]. Komposisi pembagian data pada penelitian [13], sudah mewakili keseluruhan data training dan testing. Data training digunakan untuk mencari parameter terbaik dari metode *Long Short Term Memory*. Hasil parameter terbaik tersebut akan diuji pada data testing. Jadi langkah awal yang dilakukan adalah : curah hujan dalam bentuk file CSV , seperti pada Tabel 1.

Tabel 1. Data curah hujan

No	Bulan	Tahun	Curah Hujan
1	1	2016	113.9
2	2	2016	548.3
3	3	2016	72.1
4	4	2016	35.8
...
68	10	2021	143.8
69	11	2021	423.5
70	12	2021	658

Data rata-rata curah hujan dapat digambarkan dengan bentuk grafik seperti yang ditunjukkan pada Gambar 3 di bawah ini,

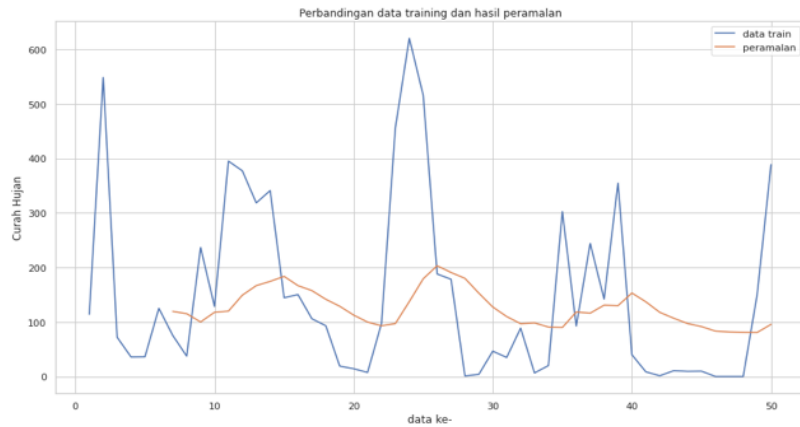


Gambar 3. Data Rata-Rata Curah Hujan Dalam Bentuk Grafik

dimana angka 0 sampai 100 merupakan rata-rata curah hujan dan pada keterangan data ke menunjukkan urutan data dari 1 sampai dengan 70, dan grafik yang berwarna biru merupakan hasil dari data rata-rata curah hujan seperti pada Gambar 3.

3.2 Perbandingan Data Testing dan Hasil Peramalan.

Hasil pengujian menggunakan data testing, terperinci menjadi tiga bagian tabel yaitu: data ke, curah hujan, dan peramalan. Gambar 4 merupakan Grafik Perbandingan Data Testing dan Hasil Peramalan. Pada gambar 4.22 garis biru menunjukkan data aktual rata-rata curah hujan kota Denpasar. Sedangkan garis *orange* menunjukkan data hasil peramalan menggunakan metode *Long Sort Term Memory* (LSTM).



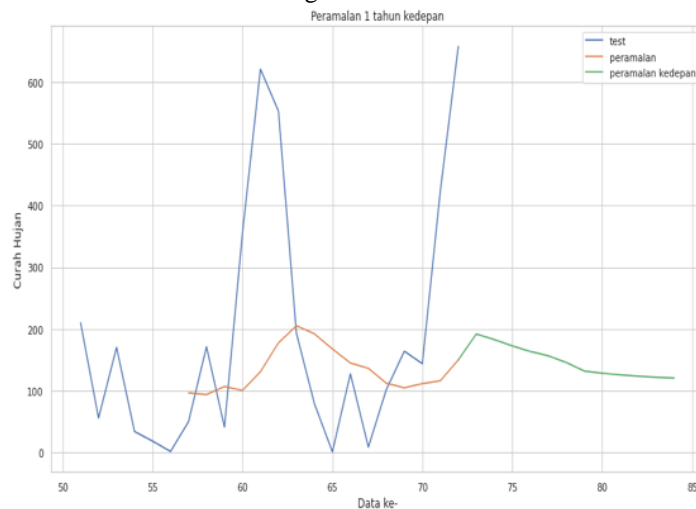
Gambar 4. Grafik Perbandingan Data Testing dan Hasil Peramalan.

3.3 Hasil Perhitungan MAE dan MAPE

Mengukur akurasi peramalan dihitung *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) pada data training dan data testing. Pada Gambar 4 menunjukkan hasil perhitungan MAE dan MAPE pada peramalan data training yaitu MAE sebesar 199.19741 dan MAPE sebesar 2.518%. sedangkan pada peramalan data testing, didapat MAE sebesar 166.56504 dan MAPE sebesar 11.89371%. Hasil peramalan pada data testing mendapatkan error yang cukup besar, hal tersebut disebabkan oleh terdapat lonjakan rata-rata curah hujan yang cukup tinggi.

3.4 Hasil Peramalan Satu Tahun kedepanya

Pada Gambar 5 menampilkan data peramalan 1 tahun kedepan dalam bentuk grafik menggunakan library matplotlib dan seaborn. Hasil dari proses peralamalan 1 tahun kedepan ditunjukkan pada Gambar 5, Perhitungan peramalan 1 tahun kedepan ini dihitung menggunakan data testing yang sebelumnya berjumlah 30% dari total data diluar data training.



Gambar 5. Grafik Hasil Peramalan satu tahun.

Pada Gambar 5, dapat dilihat garis biru merupakan data aktual rata-rata curah hujan pada Kota Denpasar, garis orange merupakan hasil peramalan dari data aktual, dan garis hijau merupakan hasil peramalan 1 tahun kedepan yang dihitung dari data testing. Pada gambar 4.29 dibawah ini data dimulai dari data ke 50 karena dari total 70 data 70% dari data digunakan untuk data training, sedangkan data uji hanya menggunakan 30% dari total data sehingga nilai awal pada data testing dimulai dari 50 sampai dengan 70.

Mengetahui hasil prediksi selama satu tahun kedepan, pihak BMKG dapat menginformasikan ke pada masyarakat prediksi curah hujan kedepannya. Antisipasi terhadap prediksi curah hujan dapat dilakukan lebih awal, diantaranya persiapan musim tanam dan panen untuk sektor pertanian dan pada sektor konstruksi, dapat mengambil waktu yang tepat saat pengerjaan kontruksi bangunan di mulai.

4. Kesimpulan

Hasil penelitian dengan melakukan prediksi curah hujan dengan menggunakan Long Short Term Memory (LSTM) menghasilkan beberapa nilai prediksi. Untuk prediksi atau peramalan curah hujan pada kota Denpasar, menunjukkan bahwa hasil prediksi yang di dapat mengasilkan nilai error atau tingkat kesalahan yang kecil. Sehingga dapat di simpulkan bahwa pada penelitian dengan arsitektur Long Short Term Memory (LSTM) menghasilkan nilai prediksi selama satu tahun. Hasil peramalan curah hujan menggunakan data pelatihan sebesar 70% dan data pengujian sebesar 30% dengan jumlah neuron optimal didapat pada jumlah neuron 30. Pengukuran akurasi metode Long Short Term Memory pada penelitian ini menggunakan metode Mean Absolute Error (MAE) dan Mean Absolute Percentage Error (MAPE), dimana pada data training, didapatkan nilai MAE sebesar 119.19741 dan MAPE sebesar 2.518%. Sedangkan pada data testing didapat MAE sebesar 166.56504 dan MAPE sebesar 11.89371%. Hasil peramalan pada data testing mendapatkan error yang cukup besar, hal tersebut disebabkan oleh terdapat lonjakan rata-rata curah hujan yang cukup tinggi pada beberapa data.

Saran yang diberikan pada penelitian ini adalah melakukan pengkajian ulang terhadap arsitektur LSTM penentuan jumlah parameter karena kemungkinan besar jumlah parameter berdampak pada hasil prediksi. Dengan demikian diperlukan hasil eksperimen atau penelitian terkait lebih lanjut mengenai beberapa jumlah dan apa saja parameter yang perlu ditambahkan untuk menghasilkan prediksi terbaik menggunakan arsitektur *Long Short Term Memory (LSTM)* pada prediksi curah hujan kota Denpasar. Jumlah data yang digunakan harus lebih banyak agar hasilnya semakin lebih baik. Diperlukan penelitian terkait penggabungan metode optimasi yang lain seperti Algoritma Genetika, *Particle Swarm Optimization*, dan Ant Colony untuk mencari jumlah neuron yang optimal.

Daftar Pustaka

- [1] S. Mopangga, "Analisis Neraca Air Daerah Aliran Sungai Bolango," *RADIAL J. Perad. Sains, Rekayasa dan Teknol.*, vol. 7, no. 2, pp. 162–171, 2020.
- [2] Rheza Adi Prihardani, "Pengertian Curah Hujan, Macam, dan Proses Terjadinya," 4 oktober 2018, 2018. .
- [3] R. Arizon, "Prediksi curah hujan menggunakan recurrent neural network rainfall prediction using recurrent neural network," 2011.
- [4] L. Wiranda and M. Sadikin, "Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma," *J. Nas. Pendidik. Tek. Inform.*, vol. 8, no. 3, pp. 184–196, 2019.
- [5] L. Yang, Q. Liu, Y. Hao, and X. Y. Zhu, "Ship Traffic Volume Forecast in Bridge Area Based on Enhanced Hybrid Radial Basis Function Neural Networks," *3rd Int. Conf. Transp. Inf. Saf. (Ictis 2015)*, pp. 38–43, 2015.
- [6] P. Sugiartawan, A. A. Jiwa Permana, and P. I. Prakoso, "Forecasting Kunjungan Wisatawan Dengan Long Short Term Memory (LSTM)," *J. Sist. Inf. dan Komput. Terap. Indones.*, vol. 1, no. 1, pp. 43–52, 2018.
- [7] F. Beaufays, H. Sak, and A. Senior, "Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling Has," *Interspeech*, no. September, pp. 338–342, 2014.
- [8] Z. C. Lipton, "A Critical Review of Recurrent Neural Networks for Sequence Learning," *CoRR*, vol. abs/1506.0, pp. 1–38, 2015.
- [9] A. Senior, H. Sak, and I. Shafran, "Context dependent phone models for LSTM RNN acoustic modelling," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2015-Augus, pp. 4585–4589, 2015.
- [10] S. G. K. Patro and K. K. Sahu, "Normalization: A Preprocessing Stage," *arXiv Prepr. arXiv1503.06462.*, p. 4, 2015.
- [11] P. Sugiartawan, R. Pulungan, and A. K. Sari, "Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 2, pp. 326–332, 2017.
- [12] M. P. Schambach and S. F. Rashid, "Stabilize sequence learning with recurrent neural networks by forced alignment," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 1270–1274, 2013.
- [13] P. Sugiartawan, R. Pulungan, and A. Kartika, "Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 2, 2017.