

Implementasi Long Short-Term Memory (LSTM) untuk Prediksi Intensitas Curah Hujan (Studi Kasus: Kabupaten Malang)

Thoriq Afa Faisal Muhammad, dan Mohammad Isa Irawan
Departemen Matematika, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: mii@its.ac.id

Abstrak—Curah hujan merupakan salah satu fenomena alam yang dianggap sebagai salah satu faktor terpenting bagi setiap orang untuk meningkatkan produktivitasnya dalam berbagai sektor usaha. Kondisi ini sangat mempengaruhi dalam pengambilan keputusan yang optimal pada aspek kehidupan dengan berbagai tujuan, salah satu contohnya adalah kegiatan manusia di sektor pertanian. Sulitnya memprediksi curah hujan dikarenakan tidak menentukannya keadaan cuaca. Pada beberapa daerah yang terlihat cerah, tidak lama kemudian dapat terjadi hujan bahkan badai. Kabupaten Malang merupakan daerah yang mempunyai iklim tropis dan juga memiliki sumber daya alam yang melimpah di sektor pertanian dan perkebunan. Pada sektor ini terdapat beberapa faktor yang memiliki pengaruh yang pada tingkat produktivitas yang mana salah satunya adalah curah hujan. Dengan dilakukannya prediksi pada curah hujan, yang bertujuan untuk meningkatkan produktivitas dan mobilitas pada aktivitas manusia. Penelitian ini membahas tentang prediksi curah hujan di Kabupaten Malang. Salah satu metode yang digunakan untuk memprediksi kondisi cuaca yaitu menggunakan *Long Short-Term Memory* (LSTM). Hasil penelitian ini diperoleh bahwa Model *Long Short-Term Memory* mempunyai performa terbaik dengan parameter yang telah ditentukan, dimana tingkat nilai error yang digunakan pada penelitian ini menggunakan RMSE dan MAE terkecil berturut-turut adalah sebesar 0.98162 dan 0.68847. Hal ini menunjukkan bahwa semakin kecil tingkat nilai error, maka semakin akurat model tersebut melakukan prediksi.

Kata Kunci—Prediksi, Curah Hujan, *Long Short-Term Memory* (LSTM).

I. PENDAHULUAN

INDONESIA dikenal negara agraris karena sebagian penduduk di Indonesia memiliki sumber mata pencaharian sebagai petani atau bercocok tanam. Indonesia juga dikenal negara yang kaya akan sumber daya alam yang melimpah, baik di darat maupun di perairan. Pada saat ini sektor pertanian dan perkebunan di Indonesia juga masih menjadi sumber yang cukup luas dan memiliki peran yang cukup penting dalam meningkatkan perekonomian maupun memenuhi kebutuhan pangan di Indonesia [1]. Pada sektor ini terdapat beberapa faktor yang memiliki pengaruh yang pada tingkat produktivitas yang mana salah satunya adalah curah hujan. Secara umum, curah hujan sendiri terbentuk dari hasil interaksi sistem udara multi skala, dan juga dipengaruhi banyak faktor lingkungan. Kondisi ini sangat berkaitan dengan kehidupan penduduk pada daerah tersebut. Oleh karenanya perubahan dinamis yang kompleks di dalam atmosfer dan butuhnya prediksi untuk curah hujan tersebut.

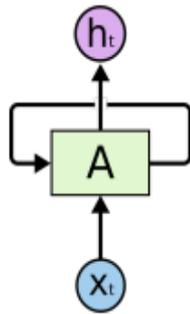
Pada masa ini keadaan cuaca di Indonesia terkadang tidak menentu. Pada beberapa daerah yang terlihat cerah, tidak lama kemudian dapat berubah menjadi hujan bahkan terjadi badai. Keadaan cuaca ini selalu diumumkan oleh Lembaga

Pemerintah Non kementerian Indonesia yaitu Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) yang bertugas di bidang meteorologi, klimatologi, dan geofisika. Menurut BMKG Indonesia, suhu udara rata-rata pada bulan Juli di Indonesia periode 1981-2010 sebesar 26,39°C. Dan pada Tahun 2020, suhu udara rata-rata pada bulan Februari sebesar 27,22°C sehingga anomali peningkatan suhu udara rata-rata sebesar 0,83°C. Kondisi cuaca sangat berkaitan dengan kegiatan manusia, karena faktor cuaca sangat mempengaruhi keterbatasan dalam kegiatan yang dilakukan oleh manusia.

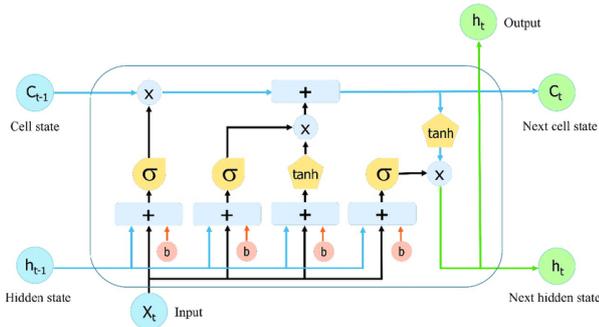
Kabupaten Malang merupakan daerah yang mempunyai iklim tropis dan juga memiliki sumber daya alam yang melimpah di sektor pertanian dan perkebunan. Pada sektor pertanian ini tingkat produksi mengalami penurunan yang mana lebih tepatnya pada produksi kopi. Kegagalan produksi panen ini disebabkan adanya cuaca yang ekstrem dan adanya curah hujan yang terlalu tinggi saat musim bunga kopi. Sehingga dengan dilakukannya prediksi pada curah hujan, yang bertujuan untuk meningkatkan produktivitas dan penjadwalan dengan menggunakan tanaman sela seperti pisang, jahe, cengkeh dan lain-lain.

Berdasarkan dari penjelasan sebelumnya, maka dibutuhkan pemilihan metode untuk memprediksi curah hujan yang dialami pada daerah tersebut. Penentuan metode yang tepat ini diperoleh dari kegiatan para peneliti atmosfer atau curah hujan yang dilakukan pada akhir-akhir ini. Dalam penelitian Jianhu Zheng, Mingfang Huang (2020) yang membahas tentang penggunaan *deep learning* untuk meramalkan arus lalu lintas melalui deret waktu. Penulis mengusulkan untuk mengembangkan model perkiraan arus lalu lintas berdasarkan LSTM [2]. Sedangkan untuk penelitian lain oleh Yuli Sudriani, Iwan Ridwansyah, Hadiid Agita Rustini (2019) yang membahas tentang prediksi dan peramalan debit air pada irigasi di Cilandir yang diperoleh dari sungai lima hari ke depan untuk membantu keputusan pembuat pada pengendalian operasi irigasi menggunakan algoritma *Deep Learning* yang melibatkan LSTM [3].

Untuk memprediksi curah hujan tersebut dapat dilakukan dengan berbagai macam cara. Salah satu cara yang bisa digunakan yaitu menggunakan *machine learning*. Dengan menggunakan beberapa metode *machine learning*, didapatkan metode terbaik yang digunakan untuk prediksi curah hujan pada *Recurrent Neural Networks* (RNN) yang mana metode tersebut merupakan salah satu bentuk arsitektur *Artificial Neural Networks* (ANN). Dari studi penelitian yang pernah dilakukan dapat disimpulkan bahwa model dari ANN mempunyai karakteristik yang membuatnya cocok dalam mengatasi kasus proses prediksi, dan secara garis besar model ANN ini juga mampu memberikan performa yang



Gambar 1. Arsitektur RNN.



Gambar 2. Arsitektur LSTM.

Tabel 1.

Ambang batas nilai intensitas hujan

Jenis Hujan	Intensitas Laju Hujan
Berawan	0 Mm/Hari
Hujan Ringan	0,5 – 20 Mm/Hari
Hujan Sedang	20 – 50 Mm/Hari
Hujan Lebat	50 – 100 Mm/Hari
Hujan Sangat Lebat	100 – 150 Mm/Hari
Hujan Ekstrem	>150 Mm / Hari

Tabel 2.

Contoh data yang tidak terdeteksi

Tanggal	Tavg
04-01-2017	24,7
05-01-2017	24,3
06-01-2017	24
07-01-2017	
08-01-2017	24,8
09-01-2017	24

memuaskan dalam proses prediksi curah hujan.

Pada penelitian akan melakukan prediksi dengan menggunakan metode model LSTM yang dapat menangani pekerjaan pada data berskala besar dan memiliki kemampuan untuk menyimpan memori dalam jangka waktu yang cukup lama. Dalam kasus ini akan di prediksi curah hujan di Kabupaten Malang yang terletak pada iklim tropis sehingga memiliki struktur data yang bervariasi. Sehingga diharapkan pada penelitian dengan model yang diusulkan dapat memprediksi curah hujan dengan akurat.

II. TINJAUAN PUSTAKA

A. Curah Hujan

Curah hujan adalah proses dimana jatuhnya air pada permukaan tanah yang mempunyai selang periode tertentu dan diukur menggunakan satuan milimeter(mm) diatas permukaan tanah horizontal tersebut. Dengan pengertian lain, curah hujan juga adalah ketinggian air dari hujan yang

Tabel 3.

Contoh data setelah di preproses

Tanggal	Tavg
04-01-2017	24,7
05-01-2017	24,3
06-01-2017	24
07-01-2017	24,4
08-01-2017	24,8
09-01-2017	24

Tabel 4.

Contoh data sebelum normalisasi

Tanggal	Tavg	RH avg	RR	Ddd x	Ff avg
04-01-2017	24,7	87	1	250	2
05-01-2017	24,3	85	1	170	1
06-01-2017	24	90	0	170	2
07-01-2017	24,8	84	5,3	190	1
08-01-2017	24,8	87	1	250	2
09-01-2017	24	85	1	170	1

Tabel 5.

Contoh data hasil normalisasi

Tanggal	Tavg	RH avg	RR	Ddd x	Ff avg
01-01-2017	0.0068	0.6315	0.8596	0.6918	0.3333
02-01-2017	0.0068	0.6052	0.8245	0.4677	0.1666
03-01-2017	0	0.4868	0.9122	0.4677	0.3333
04-01-2017	0.0365	0.6710	0.8070	0.5238	0.1666
01-01-2017	0.0068	0.6315	0.8596	0.6918	0.3333
02-01-2017	0.0068	0.6052	0.8245	0.4677	0.1666

Tabel 6.

Hasil parameter terbaik

Uji Coba	RMSE	MAE
1	1.08795	0.95256
2	2.42830	1.47452
3	0.98162	0.68847

terkumpul pada tempat tertentu. Tiap per 1 milimeter curah hujan yang berarti terdapat dalam luasan 1m² tertampung air yang memiliki ketinggian satu milimeter atau sebanyak satu liter. Dalam rentang waktu yang telah ditentukan oleh Lembaga BMKG. Kondisi tersebut dapat ditetapkan sebagai permulaan musim hujan. Terdapat ambang batas nilai yang digunakan dalam penentuan intensitas hujan sebagai berikut pada Tabel 1.

B. Artificial Intelligence (AI)

Gagasan AI sudah ada sejak tahun 1950 ketika Alan Turing pertama kali mengemukakan bahwa sebuah mesin dapat berkomunikasi dengan cukup baik untuk meyakinkan seorang manusia yang bertindak sebagai evaluator bahwa mesin tersebut juga adalah manusia. AI mewakili serangkaian kemajuan signifikan dalam teknologi, namun ini bukanlah kemajuan yang pertama, dan sepertinya tidak akan menjadi yang terakhir. Melihat kembali pada beberapa dekade terakhir, kemunculan komputer, PC, *spreadsheet*, *database relational*, konektivitas yang canggih, dan kemajuan teknologi serupa, semuanya mempengaruhi bagaimana organisasi beroperasi dan mencapai tujuannya. AI dapat melakukan hal yang sama dengan potensi untuk menjadi atau lebih *disruptive* (menyebabkan perubahan) daripada banyak kemajuan teknologi sebelumnya.

Cecerdasan Buatan (AI) adalah istilah umum yang mengacu pada teknologi yang mampu membuat mesin menjadi "cerdas." Organisasi berinvestasi dalam penelitian dan aplikasi AI untuk mengotomatisasi, meningkatkan, atau

mereplikasi kecerdasan manusia analisis dan pengambilan keputusan manusia dan profesi audit internal harus siap untuk berpartisipasi penuh dalam inisiatif organisasi dalam menerapkan AI.

Ada banyak istilah lain yang terkait dengan AI, seperti *deep learning*, *machine learning*, pengenalan gambar, pemrosesan bahasa alami, *cognitive computing* (komputasi yang mampu mengenali sesuatu), amplifikasi kecerdasan, peningkatan kognitif, peningkatan kecerdasan mesin, dan peningkatan kecerdasan.

Ciri-ciri kecerdasan buatan yaitu: (1) Mampu melakukan prediksi dan adaptasi, (2) membuat keputusan dengan dirinya sendiri, (3) belajar secara kontinu.

AI dapat dipandang sebagai kemajuan signifikan terbaru dalam rangkaian kemajuan yang telah terjadi dalam perkembangan teknologi. Apa yang menjadi perkembangan terbaru adalah kemajuan dan skalabilitas teknologi yang membuat penerapan AI dapat dipergunakan secara praktis.

C. Deep Learning

Pada masa ini Deep Learning merupakan algoritma di era Big Data [4]. Deep Learning telah menjadi pusat penelitian dan juga dilakukan pengembangan dan terobosan-terobosan baru dalam beberapa tahun terakhir. Dalam *deep learning*, terutama pada *neural networks*, telah mencapai hasil yang sangat baik dalam *search technology*, *data mining*, mesin penerjemah, pemrosesan *natural language*, *multimedia learning* dan pada bidang lainnya [5].

Pada *deep learning* terdapat banyak *layer* pemrosesan informasi tahapan dalam arsitektur hierarkis digunakan untuk *unsupervised learning* dan untuk pola klasifikasi. *Deep learning* berada di antara area penelitian jaringan saraf, pemodelan grafis, optimasi, pengenalan pola, dan pemrosesan sinyal. Dua alasan penting popularitasnya algoritma *deep learning* ini adalah biaya perangkat keras komputasi yang menurun secara signifikan dan chip yang meningkat secara drastis dalam kemampuan pemrosesan (misalnya, unit GPU) [6].

D. Recurrent Neural Network (RNN)

RNN merupakan salah satu arsitektur dari *neural networks* yang melakukan proses berulang-ulang pada inputan yang biasanya berupa data sekuensial seperti data *time-series*. Pada RNN ini berbeda dengan *neural network* tradisional dikarenakan pada setiap pemrosesan RNN yang dihasilkan tidak dipengaruhi oleh arus input tetapi juga dipengaruhi oleh keadaan internal yang mana merupakan hasil dari pemrosesan input sebelumnya. Ini berarti bahwa Ketika RNN membuat keputusan pada langkah waktu t dapat dipengaruhi terhadap keputusan yang diambil pada Langkah waktu $t - 1$. RNN sendiri sudah sangat banyak digunakan pada proses peramalan atau prediksi [7]. Arsitektur sederhana dari RNN ditunjukkan pada Gambar 1.

E. LSTM

LSTM adalah jenis modul lain dari *Recurrent Neural Network* [8]. LSTM diciptakan oleh Hochreiter and Schmidhuber (Saxeenand, 2018). Tidak seperti *single hidden layer* pada RNN, LSTM menyimpan informasi dalam unit kontrol di luar aliran normal RNN. LSTM juga memiliki struktur rantai seperti struktur RNN, perbedaannya terletak pada struktur modul pengulangannya [9]. LSTM cocok

digunakan untuk masalah yang memiliki ketergantungan jangka panjang Mengingat informasi jangka panjang adalah perilaku bawaan LSTM. Pada struktur LSTM, terdapat blok memori yang berisi tiga gerbang (*memory gate*, *forgetting gate*, and *output gate*). Di sisi lain, LSTM juga memiliki unit memori, yang mana berfungsi untuk mengontrol transfer informasi ke tahap berikutnya [10].

Berikut cara kerja arsitektur LSTM dari Gambar 2 yang dijelaskan oleh persamaan dibawah ini [9]:

1) Forget Gate

Pada *forget gate* menentukan berapa banyak informasi yang disimpan pada waktu sebelumnya atau informasi yang dihilangkan.

$$f_t = \sigma(w_f \times [h_{t-1}, x_t] + b_f) \quad (1)$$

Keterangan:

w_f : Bobot matriks *forget gate*

x_t : Data *input*

h_{t-1} : Output dari blok memori sebelumnya

b_f : Bias dari *forget gate*

σ : Fungsi sigmoid

2) Input Gate

Pada *input gate* menentukan berapa banyak informasi yang diperoleh dari x_t yang disimpan dalam *cell state* c_t :

$$i_t = \sigma(w_i \times [h_{t-1}, x_t] + b_i) \quad (2)$$

3) Output Gate

Pada *output gate* mempunyai hubungan pada hasil keluaran h_t .

$$o_t = \sigma(w_o \times [h_{t-1}, x_t] + b_o) \quad (3)$$

Pada semua proses transisi *cell* yang lama c_{t-1} ke *cell* yang baru c_t dapat dilihat pada persamaan berikut yang mana h adalah fungsi aktivasi [11].

$$\tilde{c}_t = \tanh(w_c \times [h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (5)$$

$$h_t = o_t \times \tanh(c_t) \quad (6)$$

Dari persamaan diatas dapat diketahui bahwa input LSTM tidak hanya menghasilkan output h_{t-1} dari neuron *hidden layer* pada tahap terakhir, tetapi juga mengandung nilai unit memori pada unit LSTM. LSTM dapat secara efektif menghindari kehilangan gradien, dapat mengingat informasi historis dalam jangka panjang, dan cocok untuk jangka panjang pada data *time series* yang lebih efektif.

F. Kriteria Evaluasi

Untuk mengevaluasi performa model dengan metode yang digunakan, penulis menggunakan *mean absolute error* (MAE) dan *root mean square error* (RMSE) dengan persamaan berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y'_i - y_i| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (|y'_i - y_i|)^2} \quad (8)$$

Dimana y_i adalah target output dan y'_i adalah prediksi [12].

III. METODE PENELITIAN

A. Pengumpulan Data

Bahan data yang digunakan pada penelitian ini menggunakan data yang diperoleh dari BMKG. Data yang didapat berlokasi di BMKG Meteorological Office (Stasiun Meteorologi BMKG Karangploso Malang). Variabel penelitian ini meliputi jumlah curah hujan (mm), suhu, kelembaban udara, kecepatan angin, dan tekanan udara. Pada tahap ini dilakukan seleksi pada tiap-tiap atribut yang mempengaruhi pada proses prediksi curah hujan.

B. Preprocessing Data

Pada tahap ini dilakukan dilakukan *preprocessing* data untuk memperbaiki presisi dan kinerja pada pengolahan data dan mencegah kesalahan didalam *data mining* sehingga data yang diolah dalam proses prediksi dapat berkualitas. Karena pada data yang didapat ada data yang tidak terukur mengakibatkan kegagalan pada data, maka perlunya dilakukan tahap interpolasi dari data sebelumnya dan data sesudahnya. Untuk meminimalkan *error*, dilakukan normalisasi pada data yang mana data aktual tersebut diubah menjadi dengan nilai interval $[0,1]$. Terdapat persamaan yang digunakan untuk melakukan normalisasi yaitu sebagai berikut:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (9)$$

Keterangan:

- x' : Hasil data yang dinormalisasi
- x : Data asli
- x_{min} : Nilai minimum dari keseluruhan data
- x_{max} : Nilai maksimum dari keseluruhan data

C. LSTM untuk Prediksi Curah Hujan

Pada tahap ini dilakukan pembentukan model LSTM pada data curah hujan dengan parameter-parameter yang telah ditentukan dan dilakukan perbaikan untuk kinerja LSTM dengan optimasi sehingga didapat hasil prediksi pada curah hujan.

Pada tahap ini dilakukan pemisahan data menjadi dua bagian yaitu data *training* dan data *testing*. Data *training* tersebut dipakai untuk proses pembentukan LSTM, sedangkan data *testing* digunakan untuk memvalidasi model yang telah dibuat. Pembagian data ini terdapat rasio perbandingan untuk data *training* dan data *testing* seperti contoh yaitu sebesar 80% dan 20%.

D. Proses Prediksi

Pada tahap ini proses prediksi dilakukan dengan menggunakan LSTM. Terdapat hasil-hasil dari berbagai skenario observasi pada data yang telah dilakukan pembentukan parameter-parameter akan dibandingkan dimana parameter-parameter yang berkaitan dari setiap skenario percobaan memiliki hasil peforma yang terbaik.

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada tahap ini dibutuhkan data yang digunakan dalam

penelitian tugas akhir. Pengambilan data dilakukan dengan mengunduh data dari situs website Badan Pusat Statistik yang mana berisikan data curah hujan dan diambil dalam bentuk harian. Data yang digunakan pada penelitian tugas akhir ini memiliki periode selama 5 tahun dari Januari 2017 – Februari 2022. Data yang diperoleh tersebut berisi beberapa variabel yang digunakan dalam penelitian ini yaitu jumlah curah hujan (mm), suhu, kelembaban udara, kecepatan angin dan arah angin.

B. Preprocessing Data

Pada data yang telah diambil tersebut masih terdapat beberapa nilai yang tidak terdeteksi pada website BMKG. Data yang tidak terdeteksi akan mengganggu pada saat pembentukan model serta pada saat training. Oleh karena itu diperlukan fase praproses data untuk membersihkan data-data tersebut sehingga layak digunakan dalam penelitian.

1) Data Cleaning

Data mentah yang telah diperoleh tersebut masih banyak mengandung nilai yang tidak terdeteksi, padahal data tersebut mengganggu penelitian ini. Tabel 2 merupakan contoh data-data yang mengandung nilai yang tidak terdeteksi. Contoh data yang tidak terdeteksi ditunjukkan dengan nilai yang kosong pada tabel 2 dimana diambil dengan rentang waktu dari tanggal 04-01-2017 sampai 09-01-2017.

Pada penelitian ini, untuk mengatasi data yang tidak terdeteksi tersebut dilakukan perhitungan menggunakan interpolasi dari data yang ada. Dari contoh pada tabel 2 terdapat data yang tidak yang tidak terukur diantara dua data. Ambil nilai dari kedua data tersebut dan mencari titik tengahnya (24 dan 24.8). Karena data yang terukur hanya satu maka titik tengah dicari diantara dua data tersebut. Data yang dirata-ratakan adalah data yang terdeteksi dan tidak bernilai 0. Setelah dilakukan proses penanganan data yang tidak terdeteksi, didapatkan contoh data yang ditunjukkan dengan rentang waktu dari tanggal 04-01-2017 sampai 09-01-2017 pada Tabel 3.

Kemudian apabila terdapat lebih dari satu data yang tidak terdeteksi diantara 2 data yang terukur, maka langkah yang diambil dengan cara menghitung selisih data dari nilai sebelum atau sesudah kemudian bagi dengan jarak data($data1 \rightarrow x \rightarrow x2$). Nilai selisish tersebut diambil dengan menentukan jarak nilai data sebelum(data awal) hingga data sesudah (data akhir).

Apabila terdapat 3 data atau lebih yang tidak terukur ataupun tidak ada data, maka cara yang dilakukan sama seperti yang telah dijelaskan sebelumnya. Yang perlu diperhatikan pada kasus ini adalah selisih dari data sebelum dan sesudah, serta memperhatikan jarak nilai dari kedua data tersebut. Berikut persamaan yang digunakan untuk melakukan interpolasi pada data yang tidak terdeteksi:

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) \quad (10)$$

2) Normalisasi Data

Dalam proses menuju model pada LSTM dilakukan normalisasi dataset terlebih dahulu untuk meminimalisir *error*, maka pada data tersebut dilakukan perubahan nilai dengan range interval 0 – 1 menggunakan rumus *min-max scalling*. Berikut dapat ditunjukkan contoh Tabel 4 dengan rentang waktu dari tanggal 04-01-2017 sampai 09-01-2017.

Setelah perhitungan normalisasi dihitung secara keseluruhan pada data yang akan digunakan menuju model LSTM, berikut hasil normalisasi data dapat ditampilkan pada Tabel 5 dengan mengambil rentang waktu dari tanggal 01-01-2017 sampai tanggal 02-01-2017.

C. Hasil Prediksi Menggunakan LSTM

Pada pembuatan data training merupakan proses pengenalan pola yang dilakukan dalam rangka menyesuaikan antara data input dengan data *output/target* yang diharapkan dan juga untuk meningkatkan kinerja dari model yang digunakan terhadap *data testing*. *Data training* memiliki jumlah lebih besar dibandingkan dengan *data testing* yang mempunyai maksud agar pada proses pembelajaran lebih terlatih dengan pola data dari *data training* yang nantinya akan menghasilkan suatu model dan diberikan kepada *data testing* dapat memberikan prediksi data testing yang akurat. Proses pelatihan tersebut dilakukan menggunakan metode LSTM yang menjadi model dan akan diuji performansinya terhadap *data testing*. Hasil dari proses prediksi tersebut masih berupa data normalisasi sehingga dilakukan pengembalian ke data asli/aktual dengan menggunakan proses denormalisasi. Perhitungan pada proses denormalisasi dapat dilihat di persamaan (11):

$$Y = x'(x_{max} - x_{min}) + x_{min} \quad (11)$$

Pada proses tersebut terus dilakukan perulangan hingga mendapatkan model dengan akurasi yang paling baik yang mana model tersebut dapat digunakan untuk proses prediksi.

D. Pembahasan

Pada penelitian ini dilakukan percobaan menggunakan 3 skenario. Dimana pada skenario tersebut terdapat komposisi *data training* dan *data testing* berturut-turut yaitu 60% dengan 40%, 80% dengan 20% dan 90% dengan 10%. Tiap satu skenario dilakukan percobaan dengan kombinasi dari parameter-parameter yang digunakan yaitu *epoch* dan *batch size*. Dari seluruh hasil percobaan yang telah dilakukan menggunakan kombinasi parameter dan skenario percobaan, selanjutnya akan dianalisis lebih lanjut pada model-model dengan nilai yang paling optimal. Hasil dengan menggunakan parameter-parameter terbaik dari beberapa skenario percobaan akan ditunjukkan pada Tabel 6.

Pada Tabel 6 dapat dilihat penggunaan parameter-parameter yang telah ditentukan pada model dengan tingkat nilai *error* dimana pada penelitian ini menggunakan RMSE dan MAE terkecil yang didapat berturut-turut yaitu 0.98162 dan 0.68847. Semakin kecil tingkat nilai *error* yang diperoleh, maka semakin baik model dengan parameter tersebut pada proses pembelajarannya. Hal ini membuktikan bahwa model tersebut mampu untuk memprediksi intensitas curah hujan yang diperoleh. Dengan demikian, dapat disimpulkan bahwa model *Long Short-Term Memory* memiliki performa terbaik pada uji coba 3 yang mana menggunakan kombinasi dengan *data train* dan *data testing* sebesar 90% dan 10% dengan parameter yang digunakan yaitu *learning rate* = 0.001, *epoch* = 500 dan *batch size* = 15.

E. Uji Akurasi

Berdasarkan dari pembahasan yang telah dijelaskan diatas, prediksi curah hujan tentu akan berpengaruh terhadap aktivitas dan tingkat produktivitas manusia. Apabila prediksi

yang dihasilkan semakin akurat maka hal ini akan membantu masyarakat dalam beraktivitas serta meningkatkan produktivitas khususnya pada sektor pertanian dan perkebunan. Apabila prediksi menunjukkan curah hujan yang tinggi, maka masyarakat dapat mengatur kapan waktu tanam dimulai dan bisa memaksimalkan hasil produksi panen.

V. KESIMPULAN

Berdasarkan hasil dari uji coba pada penelitian tugas akhir ini, maka kesimpulan yang dapat diambil adalah (1) Metode *Long Short-Term Memory* berhasil diterapkan dalam memprediksi intensitas curah hujan di Kabupaten Malang. Berdasarkan dari hasil pengolahan serta analisis data, dapat dibuktikan bahwa metode ini cocok dalam melakukan prediksi data sekuensial. (2) Model *Long Short-Term Memory* mempunyai performa terbaik, dimana tingkat nilai RMSE dan MAE terkecil berturut-turut adalah sebesar 0.98162 dan 0.68847. Hal ini menjelaskan bahwa prediksi intensitas curah hujan menggunakan metode *Long Short-Term Memory* dapat menghasilkan keakuratan yang baik, karena semakin kecil tingkat nilai *error* yang dihasilkan, maka semakin akurat model tersebut melakukan prediksi.

Berdasarkan hasil dan kesimpulan dari penelitian tugas akhir ini, saran yang dapat diberikan bagi penelitian selanjutnya adalah (1) Untuk penelitian selanjutnya, disarankan menggunakan jumlah data yang memiliki periode lebih panjang serta pengambilan data per jam untuk memperoleh hasil yang lebih akurat. (2) Variabel penelitian yang digunakan sebagai *input* pada model lebih bervariasi. (3) Menambah atau menggabungkan metode pengerjaan lain dalam pengerjaannya seperti menggunakan ARIMA, *Prophet*, SVM (*Support Vector Machine*) dan metode-metode lainnya untuk dijadikan metode pembanding untuk mengetahui metode mana yang paling sesuai dengan data yang digunakan.

DAFTAR PUSTAKA

- [1] Dinas Pertanian Provinsi Jawa Timur, Laporan Akuntabilitas Kinerja Instansi Pemerintah Dinas Pertanian Provinsi Jawa Timur Tahun 2013. Surabaya: Dinas Pertanian Provinsi Jawa Timur, 2013.
- [2] J. Zheng and M. Huang, "Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning," *IEEE Access*, vol. 8, pp. 82562–82570, 2020, doi: 10.1109/ACCESS.2020.2990738.
- [3] Y. Sudriani, I. Ridwansyah, and H. A. Rustini, "Long Short Term Memory (LSTM) Recurrent Neural Network (RNN) for Discharge Level Prediction and Forecast in Cimandiri River, Indonesia," in *IOP Conference Series: Earth and Environmental Science*, Jul. 2019, vol. 299, no. 1, p. 12037. doi: 10.1088/1755-1315/299/1/012037.
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [5] H. V. Ho, D. H. Nguyen, X.-H. Le, and G. Lee, "Multi-step-ahead water level forecasting for operating sluice gates in hai duong, vietnam," *Environ Monit Assess*, vol. 194, no. 6, p. 442, 2022, doi: 10.1007/s10661-022-10115-7.
- [6] M. Young, *The Technical Writers Handbook*. Mill Valley: CA: University Science, 1989.
- [7] A. Graves, S. Fernández, M. Liwicki, H. Bunke, and J. Schmidhuber, "Unconstrained On-line Handwriting Recognition with Recurrent Neural Networks," in *Neural Information Processing Systems (NIPS)*, 2007.
- [8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput*, vol. 9, pp. 1735–1780, 1997.
- [9] G. van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial intelligence review: An international survey and tutorial journal*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

- [10] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019, doi: 10.1109/TSG.2017.2753802.
- [11] Y. Li and H. Cao, "Prediction for Tourism Flow based on LSTM Neural Network," in *Procedia Computer Science*, 2018, vol. 129, pp. 277–283. doi: <https://doi.org/10.1016/j.procs.2018.03.076>.
- [12] A. H. Bukhari, M. A. Z. Raja, M. Sulaiman, S. Islam, M. Shoaib, and P. Kumam, "Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting," *IEEE Access*, vol. 8, pp. 71326–71338, 2020.