

BAB I

PENDAHULUAN

1.1 Latar Belakang

Curah hujan adalah fenomena jatuhnya air ke permukaan bumi dalam periode tertentu yang diukur dalam satuan milimeter (mm). Istilah ini juga merujuk pada volume air hujan yang turun di suatu area datar dengan asumsi tidak terjadi aliran, peresapan, atau penguapan [34]. Curah hujan memiliki peran penting dalam berbagai sektor, termasuk pertanian, industri, dan transportasi. Dalam bidang transportasi, intensitas curah hujan dapat mempengaruhi kelancaran perjalanan. Tingginya curah hujan dapat menyebabkan banjir, mengurangi jarak pandang, serta meningkatkan risiko kecelakaan. Oleh karena itu, prediksi curah hujan yang akurat menjadi hal yang sangat penting dalam membantu pengambilan keputusan dan perencanaan yang lebih baik di berbagai sektor [2].

Prediksi curah hujan merupakan proses memperkirakan jumlah curah hujan yang akan terjadi pada periode tertentu berdasarkan data historis dan faktor meteorologi yang relevan. Informasi hasil prediksi ini memiliki peran penting dalam berbagai sektor, seperti pertanian, industri, dan transportasi, terutama dalam membantu perencanaan dan pengambilan keputusan. Di Kota Surabaya, yang memiliki kepadatan penduduk tinggi serta aktivitas industri yang signifikan, prediksi curah hujan menjadi krusial dalam mendukung keberlanjutan kegiatan ekonomi dan sosial.

Dalam penelitian ini, prediksi curah hujan dilakukan dengan mempertimbangkan beberapa variabel klimatologi utama, yaitu temperatur, kelembapan, dan curah hujan. Ketiga variabel ini dipilih karena memiliki pengaruh yang signifikan terhadap pola curah hujan. Namun, dalam data cuaca sering kali terdapat data yang hilang (*missing values*) yang dapat memengaruhi akurasi prediksi. Untuk mengatasi permasalahan tersebut, digunakan metode interpolasi linear, yang berfungsi untuk menggantikan nilai yang hilang berdasarkan pola tren data yang ada. Dengan menerapkan interpolasi, diharapkan data yang digunakan dalam model menjadi lebih lengkap dan representatif terhadap kondisi aktual.

Untuk melakukan prediksi curah hujan, penelitian ini menggunakan metode *Long Short-Term Memory*, yang merupakan salah satu pendekatan berbasis deep learning yang efektif dalam menangani data deret waktu (*time series*). LSTM memiliki kemampuan untuk mempertahankan informasi jangka panjang dalam data, sehingga dapat menangkap pola yang kompleks dalam perubahan curah hujan dari waktu ke waktu. Berbagai penelitian

sebelumnya telah membuktikan bahwa metode ini mampu memberikan hasil prediksi yang lebih akurat dibandingkan metode konvensional dalam analisis data curah hujan.

Meskipun LSTM memiliki keunggulan dalam memodelkan data deret waktu, pemilihan hyperparameter yang optimal menjadi tantangan utama dalam meningkatkan performa model. *Hyperparameter* seperti *epoch*, *learning rate*, dan *time step* sangat berpengaruh terhadap akurasi prediksi. Oleh karena itu, dalam penelitian ini diterapkan *Grid Search* sebagai metode pencarian hyperparameter terbaik. Grid Search merupakan teknik pencarian sistematis yang mengevaluasi berbagai kombinasi *hyperparameter* untuk menemukan konfigurasi terbaik yang menghasilkan error terkecil, dalam hal ini diukur menggunakan *Root Mean Squared Error* (RMSE). Dengan optimasi hyperparameter melalui *Grid Search*, diharapkan performa model LSTM dapat meningkat secara signifikan.

Curah hujan dapat diprediksi dengan beberapa metode berdasarkan penelitian sebelumnya. Penelitian yang dilakukan Ni Made Meriliana Candra Devi, dkk menggunakan metode Vanilla RNN dan LSTM untuk menentukan awal musim hujan dan kemarau. Pada metode Vanilla RNN menghasilkan R^2 sebesar 0.6139, pada metode LSTM menghasilkan R^2 sebesar 0.5364. Hasil prediksi curah hujan tahun 2021 yakni Bulan Juni mulai memasuki musim kemarau dan Bulan Desember memasuki musim hujan [4].

Penelitian yang dilakukan oleh Mohanned Abduljabbar Hael & Yongsheng Yuan pada identifikasi kejadian ekstrem curah hujan menggunakan metode deteksi outlier fungsional. Hasil dari penelitian ini adalah mendapatkan hasil curah hujan di wilayah Taiz selama dua dekade terakhir yang terpengaruh kejadian ekstrem pada tahun 1999, 2004, 2005, dan 2009 [30].

Penelitian yang dilakukan oleh Brando Dharma Saputra, dkk menggunakan metode LSTM untuk menentukan prediksi curah hujan di Kabupaten Badung, Bali. Pada penelitian ini menghasilkan nilai MAPE terkecil sebesar 1.7657206683278308% dan MAE terkecil sebesar 8.97246598930908 [34].

Penelitian yang dilakukan oleh Anjali Samad, dkk mengenai prediksi curah hujan menggunakan metode *Long Short Term Memory* dan *Artificial Neural Network*. Data yang digunakan untuk melakukan pembelajaran data dari tahun 2007 hingga 2015 dan data yang digunakan untuk melakukan pengujian pada tahun 2016. Fitur yang digunakan yaitu curah hujan, kelembaban, tekanan, suhu, kecepatan angin dan arah angin. Pada penelitian ini mendapatkan RMSE terkecil berturut-turut pada 3 kota yaitu Albany 5.343, Walpole 6.280, dan Witchcliffe 7.706 [35].

Berdasarkan berbagai penelitian sebelumnya, metode *Long Short-Term Memory* terbukti efektif dalam prediksi curah hujan dan mampu menghasilkan estimasi yang akurat. Oleh karena itu, dalam penelitian ini, LSTM diterapkan untuk memprediksi curah hujan dengan mempertimbangkan variabel klimatologi utama, yaitu temperatur, kelembapan, dan curah hujan. Data yang digunakan bersumber dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dengan rentang waktu 1 Januari 2015 hingga 31 Desember 2024, yang diperoleh dari Stasiun Meteorologi Maritim Perak 1, Surabaya. Dengan menerapkan interpolasi untuk mengatasi data yang hilang dan mengoptimalkan *hyperparameter* menggunakan *Grid Search*, diharapkan prediksi curah hujan yang dihasilkan lebih akurat sehingga dapat memberikan manfaat bagi sektor-sektor yang bergantung pada informasi cuaca, khususnya di wilayah perkotaan dan industri seperti Kota Surabaya.

1.2 Perumusan Masalah

Perumusan masalah yang dibahas pada penelitian ini adalah sebagai berikut:

1.2.1 Permasalahan

Permasalahan yang terjadi dalam penelitian ini adalah implementasi metode Long Short-Term Memory (LSTM) untuk prediksi curah hujan dengan mengatasi data yang hilang menggunakan interpolasi linear dan meningkatkan akurasi model melalui optimasi hyperparameter. Serta menemukan kombinasi *hyperparameter* terbaik, yaitu epoch, learning rate, dan time step, serta menentukan rasio pembagian data yang menghasilkan nilai error terkecil dengan menggunakan metode *Grid Search*.

1.2.2 Solusi

Pada penelitian ini, metode Long Short-Term Memory (LSTM) digunakan untuk memprediksi curah hujan dengan mempertimbangkan variabel klimatologi utama, yaitu temperatur, kelembapan, dan curah hujan. Untuk mengatasi permasalahan data yang hilang (*missing values*), diterapkan Interpolasi Linear agar data lebih lengkap dan representatif terhadap kondisi aktual. Selain itu, dilakukan optimasi hyperparameter, seperti epoch, learning rate, dan time step, serta analisis rasio pembagian data terbaik menggunakan Grid Search, untuk memperoleh nilai error terkecil dan meningkatkan akurasi prediksi curah hujan.

1.2.3 Pertanyaan Penelitian

1. Bagaimana kombinasi nilai hyperparameter *epoch*, *learning rate*, dan *time step* yang menghasilkan nilai RMSE terkecil dalam penerapan metode *Long Short-Term Memory* (LSTM) untuk prediksi curah hujan setelah dilakukan interpolasi pada data yang hilang dan optimasi menggunakan *Grid Search*?
2. Bagaimana pengaruh rasio pembagian data training, validation, dan testing terhadap nilai RMSE dalam penerapan metode *Long Short-Term Memory* (LSTM) untuk prediksi curah hujan setelah dilakukan interpolasi pada data yang hilang dan optimasi menggunakan *Grid Search*?

1.3 Tujuan dan Manfaat

Tujuan dan manfaat penelitian ini adalah:

1.3.1. Tujuan Penelitian

1. Mengetahui kombinasi hyperparameter terbaik *epoch*, *learning rate*, dan *time step* yang menghasilkan nilai RMSE terkecil dalam prediksi curah hujan menggunakan metode *Long Short-Term Memory* setelah dilakukan interpolasi Linear untuk mengatasi data yang hilang serta optimasi menggunakan *Grid Search*.
2. Mengetahui rasio pembagian data terbaik yang menghasilkan nilai RMSE terkecil dalam prediksi curah hujan menggunakan metode *Long Short-Term Memory* setelah dilakukan interpolasi Linear untuk mengatasi data yang hilang serta optimasi menggunakan *Grid Search*.

1.3.2. Manfaat Penelitian

1.3.2.1. Manfaat Bagi Peneliti

Manfaat dari penelitian ini adalah memberikan manfaat bagi peneliti dalam mengetahui kombinasi *hyperparameter* terbaik, yaitu *epoch*, *learning rate*, dan *time step*, serta rasio pembagian data yang menghasilkan nilai RMSE terkecil dalam prediksi curah hujan. Penelitian ini juga membantu dalam memahami efektivitas metode *Long Short-Term Memory* setelah dilakukan interpolasi Linear untuk mengatasi data yang hilang serta optimasi menggunakan *Grid Search* dalam meningkatkan akurasi prediksi curah hujan.

1.4 Batasan Masalah

Batasan – batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) melalui situs <https://dataonline.bmkg.go.id>, dengan sumber data dari Stasiun Meteorologi Perak 1, Surabaya. Data yang digunakan berupa time-series dalam format .xlsx, dengan rentang waktu 1 Januari 2015 hingga 31 Desember 2024, yang terdiri dari 3.653 data harian.
2. Untuk menangani *missing values* dalam data, digunakan metode *Interpolasi Linear*, yang bertujuan menggantikan nilai yang hilang agar data lebih representatif dan mendukung peningkatan akurasi model prediksi.
3. Metode yang digunakan dalam penelitian ini untuk prediksi curah hujan adalah *Long Short-Term Memory*.
4. Variabel yang digunakan dalam penelitian ini adalah Temperatur, kelembapan, dan curah hujan.
5. Aplikasi menggunakan Bahasa pemrograman python dan menggunakan streamlit.
6. Optimasi *hyperparameter* dilakukan menggunakan *Grid Search* untuk menentukan kombinasi terbaik dari *epoch*, *learning rate*, dan *time step*, guna meningkatkan akurasi prediksi.
7. Evaluasi performa model dilakukan dengan menghitung nilai *error* menggunakan *Root Mean Squared Error* (RMSE) sebagai metrik utama dalam mengukur tingkat akurasi prediksi.

1.5 Sistematika Proposal

Sistematika penulisan Proposal ini terdiri dari beberapa bab, sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini membahas mengenai latar belakang, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika proposal.

BAB II KAJIAN PUSTAKA

Pada bab ini membahas mengenai teori-teori yang terkait dengan penelitian yang sedang dilaksanakan seperti tentang prediksi curah hujan, normalisasi data, *Long Short Term Memory*, evaluasi RMSE, serta penelitian terkait.

BAB III METODE USULAN

Pada bab ini membahas mengenai rancangan sistem prediksi curah hujan, dataset yang digunakan, dan skenario uji coba.