

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kota Surabaya merupakan salah satu kota besar di Indonesia yang sering mengalami curah hujan tinggi, terutama pada musim hujan. Curah hujan yang tidak menentu seringkali menyebabkan genangan air dan banjir di berbagai wilayah perkotaan, mengakibatkan gangguan aktivitas ekonomi dan sosial masyarakat. Kondisi ini diperparah dengan sistem drainase yang kurang optimal, sehingga volume air yang melimpah tidak dapat diatasi secara cepat. Selain itu, perubahan iklim global telah berdampak pada pola cuaca di Indonesia, termasuk di Surabaya, yang menjadikan prediksi cuaca semakin sulit dan penuh ketidakpastian. Oleh karena itu, Prediksi curah hujan menjadi fokus yang sangat penting karena memiliki dampak signifikan pada berbagai aspek kehidupan (Abdul Rohman dkk, 2024).

Prediksi curah hujan yang akurat memiliki peran penting dalam perencanaan kota dan mitigasi risiko bencana, terutama di daerah padat penduduk yang rentan terkena banjir. Informasi yang tepat mengenai curah hujan dapat membantu mengurangi dampak buruk dari banjir dan memberikan panduan bagi pengambilan keputusan yang lebih baik. Dalam upaya meningkatkan akurasi prediksi, berbagai metode canggih mulai diterapkan, termasuk teknologi deep learning. Salah satu metode *deep learning* yang banyak digunakan untuk analisis data cuaca adalah *Long Short-Term Memory* (LSTM).

Long Short-Term Memory (LSTM) merupakan pengembangan dari RNN yang memiliki kemampuan bagus dalam menangani interaksi nonlinear yang tinggi dan dapat mengatasi masalah gradien terutama pada kasus prediksi curah hujan (Jiangwei Zhang dkk, 2021). *Long Short-Term Memory* adalah salah satu algoritma perhitungan data time series dengan data yang digunakan seperti curah hujan, algoritma ini cukup terbukti handal dalam memprediksi beberapa tahun terakhir ini (Jamilatul Badriyah dkk, 2022). Pemilihan menggunakan metode atau algoritma LSTM (*Long Short Term Memory*) merujuk kepada penelitian terdahulu yang telah melakukan perbandingan atau keakuratan dengan menggunakan metode *Long Short-Term Memory* (LSTM).

Muhammad Rizki dalam penelitiannya berhasil memprediksi curah hujan di Kota Malang dengan akurasi diatas 85%, menunjukkan bahwa metode Long Short Term Memory (LSTM) mampu memberikan prediksi yang cepat dan akurat. Keberhasilan metode ini dalam menangkap pola data historis curah hujan membuktikan potensinya sebagai acuan dalam pengembangan model prediksi cuaca yang lebih efektif. Efisiensi ini sangat relevan, terutama untuk kota atau daerah dengan karakteristik cuaca yang serupa. Temuan ini menjadi kontribusi penting dalam meningkatkan kualitas prediksi cuaca (Muhammad Rizki dkk, 2020).

Riza Farikhul Firdaus dalam penelitiannya berhasil memprediksi curah hujan di Kota Bandung menggunakan metode *Long Short Term Memory* (LSTM). Penelitian tersebut menghasilkan nilai akurasi tertinggi dengan *Train Score* RMSE sebesar 12,24 dan *Test Score* RMSE sebesar 8,86. Nilai Test Score sebesar 8,86 menunjukkan bahwa hasil prediksi memiliki tingkat akurasi yang baik. Semakin kecil nilai *Test Score*, semakin tinggi akurasi prediksi yang dihasilkan (Riza Farikhul Firdaus dkk, 2022).

M. Devid Alam Carnegie dalam penelitiannya menganalisis dan memprediksi parameter cuaca dengan membandingkan dua algoritma, yaitu *Long Short Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU). Penelitian tersebut menunjukkan bahwa model LSTM 1 dengan pembagian dataset 7:3 memiliki performa terbaik dalam memprediksi curah hujan. Hasil evaluasi model mencatat nilai RMSE sebesar 16,81, MSE sebesar 282,55, dan MAD sebesar 10,43. Temuan ini menegaskan keunggulan LSTM dalam menghasilkan prediksi yang lebih akurat dibandingkan dengan GRU (M. Devid Alam Carnegie dkk, 2020).

Selain itu, LSTM dinilai cocok untuk melakukan prediksi karena dapat melakukan proses dengan bentuk data time series dengan melibatkan variabel variabel yang menjadi faktor yang dapat mempengaruhi terjadinya hujan. Namun, faktor yang mempengaruhi curah hujan sangat kompleks, sehingga reduksi dimensi sangat diperlukan agar model bekerja secara optimal (Jiangwei Zhang dkk, 2021), salah satunya dengan melibatkan *Principal Component Analysis* (PCA).

Principal Component Analysis (PCA) merupakan metode *feature extraction* yang digunakan untuk mereduksi dengan mengompres atribut yang jumlahnya banyak menjadi lebih sedikit (R. Pujiyanto dkk, 2019). Dalam penelitian yang dilakukan oleh Bashir, PCA digunakan untuk mengidentifikasi variabel-variabel

yang paling berpengaruh terhadap kerentanan Akuifer Varamin di Iran. Dari 14 variabel yang dianalisis, PCA berhasil menyeleksi 7 variabel utama (B. Rahmani dkk, 2021). Selain itu, penelitian oleh Tripathi juga memanfaatkan PCA dalam pemilihan parameter kualitas air untuk Sungai Gangga di India, menghasilkan 9 parameter utama dari total 28 parameter yang dianalisis (M. Tripathi dkk, 2019).

Dari beberapa penelitian serta analisis yang dilakukan, dapat disimpulkan bahwa LSTM dengan kemampuannya sangat cocok untuk memprediksi intensitas curah hujan pada data time series dan PCA dapat mempercepat pelatihan neural network, sehingga penulis memutuskan untuk meneliti terkait “Prediksi Intensitas Curah Hujan di Kota Surabaya Menggunakan *Principal Component Analysis* (PCA) dan *Long Short-Term Memory* (LSTM)”. Data yang akan digunakan untuk penelitian ini berasal dari BMKG (Badan Meteorologi, Klimatologi, dan Geofisika).

Saat ini, prediksi curah hujan tidak hanya dilakukan oleh lembaga resmi seperti BMKG, tetapi juga tersedia melalui berbagai platform digital. Aplikasi cuaca di perangkat Android, seperti *Google Weather*, biasanya memperoleh data dari penyedia seperti *The Weather Channel* atau *AccuWeather*. Di sisi lain, *Apple Weather* pada perangkat iOS menggunakan data dari *Dark Sky*, yang dikenal menyediakan informasi cuaca hiper-lokal. Selain itu, aplikasi lain seperti *Weather Underground*, *Windy*, *Meteoblue*, dan *Tomorrow.io* juga memanfaatkan kombinasi data radar, satelit, sensor cuaca pribadi, serta model numerik global seperti ECMWF dan GFS.

Namun, banyak dari *platform* tersebut tidak mengungkapkan metode atau sumber data secara rinci, sehingga sulit digunakan dalam penelitian ilmiah. Oleh karena itu, penelitian ini menggunakan data dari BMKG (Badan Meteorologi, Klimatologi, dan Geofisika), lembaga resmi pemerintah Indonesia yang menyediakan data cuaca terstandarisasi dan dapat dipertanggungjawabkan. Data yang digunakan mencakup variabel suhu rata-rata, suhu maksimum, suhu minimum, kelembapan, tekanan udara, kecepatan angin rata-rata, kecepatan angin maksimum, penyinaran matahari, dan arah mata angin di Kota Surabaya, yang kemudian diekstraksi dengan metode PCA. Hasil dari PCA kemudian digunakan sebagai input untuk memprediksi curah hujan harian dengan metode LSTM. Selanjutnya, hasil prediksi akan dievaluasi menggunakan MSE, MAE, dan RSME. Harapan penulis, Dengan menggunakan metode ini, prediksi curah hujan di Kota Surabaya dapat

dilakukan lebih adaptif terhadap perubahan pola cuaca yang dinamis dan sulit diprediksi.

1.2.Rumusan Masalah

Uraian latar belakang yang telah ditulis memiliki beberapa permasalahan yang perlu dipecahkan pada penelitian yang dilakukan, yakni:

1. Bagaimana hasil proses reduksi variabel dari data historis curah hujan di Kota Surabaya dengan *Principal Component Analysis* (PCA)?
2. Bagaimana performa prediksi intensitas curah hujan harian di Kota Surabaya menggunakan metode LSTM dengan variabel yang diekstraksi oleh PCA?
3. Bagaimana hasil perbandingan akurasi prediksi intensitas curah hujan di Kota Surabaya pada tiga skenario PCA menggunakan model LSTM?

1.3.Tujuan Penelitian

Berdasarkan perumusan masalah yang telah ditetapkan, pemodelan sistem prediksi intensitas curah hujan di Kota Surabaya menggunakan *Principal Component Analysis* (PCA) Dan *Long Short-Term Memory* (LSTM) mempunyai tujuan sebagai berikut:

1. Mengoptimalkan data historis curah hujan di Kota Surabaya dengan menggunakan *Principal Component Analysis* (PCA) untuk menghasilkan variabel yang relevan dalam prediksi curah hujan.
2. Mengetahui performa metode *Long Short-Term Memory* (LSTM) dalam memprediksi intensitas curah hujan harian di Kota Surabaya dengan menggunakan variabel yang diekstraksi melalui *Principal Component Analysis* (PCA)
3. Membandingkan akurasi prediksi curah hujan di Kota Surabaya berdasarkan tiga skenario PCA untuk menentukan konfigurasi terbaik antara akurasi dan efisiensi model.

1.4 Manfaat Penelitian

Berdasarkan penelitian ini, ada beberapa manfaat yang diperoleh yaitu:

1. Bagi Peneliti, Penelitian ini memberi penulis pemahaman mendalam tentang penerapan *Principal Component Analysis* (PCA) dan *Long Short-Term Memory* (LSTM) dalam menganalisis data cuaca berbasis time series, serta memperkuat keterampilan analisis data dengan mereduksi dimensi melalui

PCA dan mengevaluasi akurasi model prediksi menggunakan indikator evaluasi seperti MSE, MAE, dan RMSE. Penelitian ini juga menambah kontribusi ilmiah penulis dalam bidang prediksi cuaca, yang bermanfaat sebagai referensi untuk penelitian lanjutan atau pengembangan sistem prediksi cuaca di wilayah lain.

2. Bagi Masyarakat Umum, hasil prediksi curah hujan yang akurat dapat membantu masyarakat mengantisipasi risiko bencana banjir, terutama di daerah rentan di Surabaya. Masyarakat dapat lebih siap menghadapi cuaca ekstrem, merencanakan aktivitas dengan baik, dan mengurangi gangguan pada kehidupan sehari-hari. Selain itu, data prediksi ini juga mendukung pemerintah daerah dalam memperbaiki infrastruktur kota seperti sistem drainase, untuk menekan dampak banjir pada masyarakat perkotaan.
3. Bagi Pembaca, Penelitian ini menjadi sumber pengetahuan bagi pembaca tentang penggunaan PCA dan LSTM dalam prediksi cuaca, serta menambah literatur bagi mereka yang tertarik pada data science atau meteorologi. Selain sebagai inspirasi bagi penelitian serupa, pembaca juga mendapatkan wawasan tentang tantangan dalam prediksi cuaca dan pentingnya prediksi yang akurat untuk mendukung kehidupan masyarakat dan adaptasi terhadap perubahan iklim.

1.5 Batasan Masalah

Penelitian ini memiliki beberapa batasan yang harus diperhatikan, yaitu:

- a. **Lokasi Penelitian:** Penelitian hanya dilakukan untuk memprediksi curah hujan di Kota Surabaya. Hasil prediksi belum tentu dapat diaplikasikan di daerah lain tanpa penyesuaian.
- b. **Variabel Meteorologi Terbatas:** Variabel yang digunakan dalam prediksi hanya mencakup suhu rata – rata, suhu maksimal, suhu minimum, kelembapan rata – rata, tekanan, kecepatan angin rata – rata, kecepatan angin terbesar, penyinaran matahari, dan arah mata angin.
- c. **Data Historis dari BMKG:** Model prediksi hanya menggunakan data curah hujan historis dan variabel meteorologi dari BMKG Perak I tahun 2018-2024
- d. **Metode yang Digunakan:** Penelitian ini menggunakan metode *Principal Component Analysis* (PCA) untuk reduksi dimensi dan *Long Short-Term*

Memory (LSTM) untuk pemodelan prediksi curah hujan. Hasil yang diperoleh terbatas pada efektivitas kedua metode ini dan belum dibandingkan dengan metode lain.