

BAB I

PENDAHULUAN

1.1. Latar Belakang

Deep learning merupakan teknik berbasis jaringan saraf tiruan yang telah menjadi salah satu metode populer dalam implementasi *machine learning* dalam beberapa tahun terakhir [1]. *Deep learning* hadir sebagai solusi untuk mengatasi keterbatasan *machine learning* tradisional dengan menawarkan pendekatan yang lebih otomatis, fleksibel, dan mampu menangani data yang lebih kompleks [2]. Teknik ini tidak terbatas pada satu bidang saja, melainkan dirancang sebagai metode pembelajaran yang serbaguna untuk menyelesaikan berbagai masalah di berbagai sektor. Popularitas *deep learning* mulai meningkat sejak tahun 2006, dengan memanfaatkan pendekatan *deep architecture of learning* atau *hierarchical learning*. *Learning* atau pembelajaran dalam hal ini adalah sebuah prosedur yang berisi proses estimasi parameter-parameter suatu model sehingga model yang dikembangkan dapat menyelesaikan suatu tugas atau permasalahan tertentu [1]. Teknologi *deep learning* melibatkan beberapa lapisan atau *layer* yang terletak di antara lapisan *input* dan lapisan *output*. Arsitektur ini memungkinkan pemrosesan nonlinier melalui berbagai tahap, yang hasilnya dapat dimanfaatkan untuk pembelajaran *fitur* (*feature learning*) dan klasifikasi pola (*pattern classification*) [1]. Salah satu arsitektur *deep learning* yang dapat digunakan untuk pemrosesan nonlinier [3] adalah *Long Short-Term Memory* (LSTM). LSTM sendiri berfokus untuk memodelkan data deret waktu (*time series*) yang dirancang untuk menangani ketergantungan jangka panjang serta mampu mempertahankan informasi penting dari waktu ke waktu dalam proses pelatihan [4].

Salah satu contoh penerapan LSTM adalah prediksi curah hujan dasarian untuk menentukan awal musim hujan dan musim kemarau seperti yang dilakukan oleh Devi dkk yang menggunakan data Stasiun Meteorologi Kelas II Gusti Ngurah Rai dari tahun 2015-2020 [5]. Penelitian ini didasari oleh perubahan iklim yang memengaruhi pertanian, sehingga diperlukan prediksi curah hujan yang akurat. Metode yang digunakan adalah *Vanilla Recurrent Neural Network* (*Vanilla RNN*) dan *Long Short-Term Memory* (LSTM), yang keduanya merupakan jenis jaringan saraf berulang yang cocok untuk data sekuensial. Selain itu, seleksi *fitur* dengan metode *Backward Elimination* diterapkan

untuk meningkatkan akurasi prediksi. Dari beberapa *fitur* cuaca yang digunakan, hasil seleksi menunjukkan bahwa kelembaban udara, tekanan udara, dan jarak pandang adalah *fitur* yang paling relevan. Hasil akhir menunjukkan bahwa *Vanilla* RNN dengan seleksi *fitur* memberikan performa terbaik dengan nilai *R-Squared* sebesar 0,6139 dan RMSE sebesar 28,4308 [5]. Namun, penelitian ini memiliki keterbatasan data, yaitu hanya lima tahun yang dapat menyebabkan model kurang mampu menangkap pola jangka panjang dan tren perubahan iklim.

Selain prediksi musim hujan, LSTM juga diterapkan dalam memprediksi variabel cuaca lainnya, seperti iradiasi matahari yang dilakukan oleh M. V. Nugroho dkk pada tahun 2024 di Jawa-Bali menggunakan metode *Stacked Long Short-Term Memory* (LSTM) [6]. Penelitian ini memanfaatkan data historis yang diperoleh dari pengukuran langsung. Penelitian ini juga membandingkan dua metode lain yaitu ARIMA dan *Recurrent Neural Network* (RNN). Hasil dari penelitian ini menunjukkan bahwa model LSTM yang diusulkan mengungguli ARIMA dan RNN, dengan nilai RMSE sebesar 25,56 W/m², MAPE sebesar 7,27%, dan R² sebesar 0,99, yang menandakan akurasi ramalan yang sangat tinggi [6]. Model *Stacked* LSTM menunjukkan performa yang lebih baik pada musim kemarau dibandingkan musim hujan, dengan MAPE terendah sebesar 13,99% pada musim kemarau, yang dikategorikan sebagai ramalan yang baik, dan MAPE tertinggi sebesar 34,04% pada musim hujan, yang dikategorikan sebagai ramalan yang wajar. Hasil ini menunjukkan bahwa model LSTM yang diusulkan merupakan pendekatan yang menjanjikan dan unggul untuk ramalan iradiasi matahari jangka pendek di Jawa-Bali [6].

Stacked LSTM juga telah digunakan dalam bidang meteorologi untuk memprediksi kecepatan angin selama episode topan, menunjukkan kemampuan model ini dalam menangani berbagai jenis data *time series*. Penelitian yang dilakukan oleh Wei pada tahun 2020 ini dilakukan di wilayah timur laut Taiwan, yang sering terdampak oleh angin kencang dan hujan deras akibat topan [7]. Untuk meningkatkan akurasi prediksi kecepatan angin, penelitian ini menggunakan teknik *machine learning* canggih dengan menerapkan nilai simulasi dari model *Weather Research and Forecasting* (WRF) sebagai *input* variabel dan mengadopsi jaringan saraf tiruan yang lebih dalam, seperti *Multilayer Perceptron* (MLP), *Deep Recurrent Neural Networks* (DRNN), dan *Stacked* LSTM. Metode ini berbeda dalam kapasitas memorinya, dengan MLP tidak memiliki kapasitas

memori, sedangkan DRNN dan *Stacked LSTM* memiliki kapasitas memori yang memungkinkan analisis data *time-series* [7]. Hasil penelitian menunjukkan bahwa model dengan kapasitas memori DRNN dan *Stacked LSTM* lebih akurat dibandingkan model tanpa kapasitas memori MLP, dengan *Stacked LSTM* memberikan hasil yang sedikit lebih akurat dibandingkan DRNN. Selain itu, penggunaan nilai simulasi numerik dalam model prediksi angin mengurangi tingkat kesalahan prediksi sekitar 30%, sehingga meningkatkan akurasi prediksi secara keseluruhan [7]. Diluar bidang meteorologi dan klimatologi, teknologi *deep learning* seperti *stacked LSTM* diterapkan dalam berbagai bidang lain seperti yang dilakukan oleh Ma pada tahun 2022 dalam memprediksi rekaman dan tren variasi data kontrol kualitas (QC) dari mesin radioterapi menggunakan pengembangan model jaringan syaraf tiruan berbasis *stacked LSTM* [8]. Data yang digunakan terdiri dari 24 item QC yang dikumpulkan harian selama 3 tahun, dengan total 867 rekaman yang digunakan untuk memprediksi data 5 hari ke depan. Selain memprediksi, penelitian ini juga membandingkan metode LSTM dengan model ARIMA pada dataset yang sama untuk mengevaluasi akurasi. Akurasi model diukur menggunakan *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan Koefisien Determinasi (R^2). Untuk menguji kekuatan model, data dari empat item QC untuk mesin radioterapi lainnya digunakan sebagai uji coba tanpa mengubah *hyperparameter* model LSTM dan ARIMA [8]. Hasil penelitian menunjukkan bahwa model *Stacked LSTM* lebih unggul dibandingkan ARIMA, dengan nilai MAE rata-rata sebesar 0.013, RMSE 0.020, dan R^2 0.853, dibandingkan dengan nilai MAE 0.021, RMSE 0.030, dan R^2 0.618 untuk ARIMA pada 24 item QC. Selain itu, model LSTM juga menunjukkan performa yang lebih baik dalam mengolah data dari mesin radioterapi lain. Hal ini menunjukkan bahwa model *Stacked LSTM* dapat memprediksi rekaman dan tren QC dengan akurasi yang tinggi dan robust, serta mampu membantu dalam memprediksi kegagalan mesin untuk pemeliharaan dini dan pengurangan *downtime* yang tidak terjadwal [8].

Keempat penelitian tersebut memanfaatkan *deep learning* untuk memprediksi berbagai fenomena yang berbeda. Tiga di antaranya secara khusus menerapkan model *stacked LSTM* untuk meramalkan iradiasi matahari, kecepatan angin, serta variasi data kontrol kualitas pada mesin. Namun, data yang digunakan dalam masing-masing penelitian masih terbatas dalam durasi waktu. Meskipun hasil menunjukkan performa dengan tingkat akurasi yang cukup baik, model *stacked LSTM* memiliki kemampuan

yang lebih besar jika diterapkan pada data yang lebih panjang dan lebih banyak. Pengujian model dalam jangka waktu yang lebih panjang dapat memperkuat validitas prediksi, menguji keandalan model dalam skenario jangka panjang, dan menunjukkan apakah *stacked* LSTM tetap konsisten dalam menghasilkan prediksi akurat seiring dengan bertambahnya data. Hal ini penting untuk memberikan bukti yang lebih kuat tentang kemampuan *stacked* LSTM dalam memproses data *time series* yang lebih kompleks dan luas.

Melihat potensi besar dari *deep learning*, khususnya model *stacked* LSTM yang telah terbukti dalam berbagai aplikasi sebelumnya, pendekatan ini dapat menjadi salah satu alternatif untuk diterapkan dalam konteks iklim di Indonesia saat ini. Kondisi perubahan iklim yang semakin nyata sejalan dengan laporan terbaru Badan Riset dan Inovasi Nasional (BRIN) yang mengungkap peningkatan signifikan dalam kejadian kekeringan dan hujan ekstrem di berbagai wilayah. Dalam Siaran Pers BRIN pada tanggal 30 Januari 2024, Erma Yulihastin, Peneliti Ahli Utama Klimatologi dan Perubahan Iklim dari Pusat Riset Iklim dan Atmosfer BRIN, mengungkapkan bahwa frekuensi hujan ekstrem dan kekeringan telah meningkat secara drastis. Di Pulau Jawa, banyak wilayah, terutama pantura Jawa Timur, diperkirakan akan menghadapi suhu yang lebih tinggi dari suhu maksimum sebelumnya dan suhu yang lebih rendah dari suhu minimum sebelumnya [9]. Pola curah hujan juga menunjukkan perubahan signifikan seperti jumlah hari tanpa hujan berturut-turut juga meningkat sepanjang 21-30 hari atau lebih panjang yang mayoritas terjadi di wilayah Jawa, Nusa Tenggara, dan Bali yang dapat mempengaruhi pasokan air dan pertanian [10]. Selain itu, frekuensi kejadian hujan lebat dengan intensitas sangat tinggi berpotensi meningkatkan risiko bencana, seperti banjir dan kerusakan infrastruktur. Perubahan yang semakin tidak menentu ini menunjukkan bahwa karakteristik musim di Indonesia pun turut mengalami pergeseran.

Indonesia sendiri merupakan negara dengan iklim tropis yang hanya mengalami dua musim, yakni musim hujan dan musim kemarau. Umumnya, musim hujan di Indonesia terjadi antara bulan Oktober hingga Maret, sedangkan musim kemarau berlangsung dari April hingga September [11]. Meski musim berganti secara periodik, pergeseran musim dapat terjadi karena perubahan iklim dapat memengaruhi pola dan karakteristik musim di suatu wilayah, dengan dampak seperti memperpanjang, mempersingkat, atau mengubah intensitas musim, serta meningkatkan frekuensi cuaca

ekstrem. Kepala Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), Dwikorita Karnawati, menyatakan bahwa dampak perubahan iklim yang terjadi di bumi semakin memprihatinkan. Dalam Rapat Nasional Prediksi Musim Kemarau 2024 pada 9 Februari 2024, beliau juga menjelaskan bahwa kenaikan konsentrasi gas rumah kaca di atmosfer diakibatkan oleh pembakaran bahan bakar fosil, penggundulan hutan, serta praktik industri yang tidak ramah lingkungan. Faktor-faktor ini telah mendorong perubahan iklim dengan laju yang belum pernah terjadi sebelumnya. Ia juga menegaskan bahwa perubahan iklim global bukanlah hoaks atau sekadar prediksi, melainkan kenyataan yang sudah dihadapi oleh miliaran orang di seluruh dunia [12].

Dari permasalahan yang sudah dijelaskan paragraf sebelumnya, pemahaman yang lebih mendalam terhadap dinamika musiman menjadi sangat penting untuk mendukung ketahanan sektor yang bergantung pada cuaca atau musim seperti pertanian, infrastruktur, dan sumber daya air. Fenomena ini tidak bisa dianggap sepele karena dampaknya sudah terasa luas. Dalam menghadapi tantangan tersebut, prediksi berbasis data masa lalu dan masa kini menjadi alat penting untuk memperkirakan kejadian yang paling mungkin terjadi di masa depan dengan tujuan meminimalkan kesalahan dan menghasilkan estimasi yang sedekat mungkin dengan kenyataan, meskipun tidak selalu memberikan jawaban pasti [13].

Tujuan tersebut sejalan dengan penerapan model *stacked* LSTM dalam meningkatkan akurasi model sistem deteksi intrusi (*Intrusion Detection System*) yang dilakukan oleh Alsyaibani dkk pada tahun 2022 bersama model *deep learning* lainnya seperti *Gated Recurrent Unit* (GRU) [14]. Penelitian ini juga mengevaluasi dampak penambahan lapisan *bidirectional* pada kinerja model deteksi intrusi. Hasil penelitian menunjukkan bahwa metode *stacked* mencapai akurasi 98,1087% setelah 100 iterasi pelatihan, yang sedikit lebih tinggi dibandingkan LSTM, GRU, *Bidirectional* LSTM, dan *Bidirectional* GRU. LSTM menunjukkan performa terbaiknya dengan fungsi aktivasi Tanh, sementara GRU dan *Bidirectional* GRU mencapai performa terbaik dengan fungsi aktivasi Lrelu dan Prelu [14]. Meskipun mendapatkan akurasi yang cukup baik, penelitian ini menghadapi beberapa keterbatasan, termasuk masalah keterbatasan data yang dapat mengurangi kemampuan model untuk generalisasi, terutama jika dataset tidak mencakup variasi serangan yang cukup. Selain itu, penanganan nilai *null* hanya dilakukan dengan menghapus baris yang mengandung nilai tersebut, yang berisiko kehilangan informasi

berharga dan dapat menghasilkan bias dalam model. Penelitian ini juga tidak mencantumkan evaluasi metrik yang jelas guna menilai performa model secara objektif.

Berbeda dengan pendekatan *deep learning*, metode prediksi yang digunakan oleh BMKG sebelumnya adalah ARIMA (*AutoRegressive Integrated Moving Average*), yang berbasis pada pendekatan linear [15]. Sedangkan data meteorologi, seperti curah hujan sangat dinamis dimana ARIMA memiliki keterbatasan dalam menangkap pola nonlinear sehingga dapat memberikan hasil yang kurang maksimal. Selain itu, BMKG juga mengandalkan metode *Numerical Weather Prediction* (NWP) berbasis model deterministik, seperti WRF, GFS, dan ECMWF [16]. Model-model ini bekerja dengan menjalankan simulasi atmosfer secara detail menggunakan satu set kondisi awal, menghasilkan satu hasil prediksi tunggal yang bersifat presisi tinggi. Meskipun akurat untuk jangka pendek, pendekatan deterministik memiliki keterbatasan dalam menangkap ketidakpastian kondisi atmosfer, terutama untuk prediksi jangka menengah hingga panjang [17]. Di sisi lain, penelitian ini mengusulkan pendekatan *Stacked LSTM*, yang merupakan bagian dari teknik ensemble berbasis pembelajaran mesin. Dengan melatih beberapa lapisan memori jangka pendek, model ini mampu mengenali pola kompleks dalam data deret waktu dan menghasilkan prediksi berdasarkan lebih dari satu kemungkinan. Pendekatan ini penting dikembangkan karena dapat mengurangi sensitivitas terhadap kesalahan awal, memperbaiki ketepatan prediksi untuk jangka waktu yang lebih panjang, serta memberikan alternatif prediksi yang lebih adaptif terhadap dinamika atmosfer tropis seperti yang terjadi di wilayah Indonesia.

Penelitian ini berkontribusi pada pengembangan metode prediksi cuaca yang lebih andal dan informatif, khususnya dalam mendukung perencanaan sektor-sektor sensitif iklim seperti pertanian dan pengelolaan sumber daya air. Dipilihnya model *stacked LSTM* juga karena berdasarkan beberapa penelitian, terbukti dalam lebih unggul dibanding ARIMA [18], [19], [20]. Pendekatan *Stacked LSTM* yang terdiri dari beberapa lapisan LSTM yang disusun secara bertingkat digunakan untuk memprediksi periode musim merupakan inovasi potensial dan efisien yang belum pernah diterapkan sebelumnya di wilayah Jawa Timur. Selain itu, dalam penelitian ini juga ingin diketahui apakah dengan menambah jumlah lapisan LSTM dapat meningkatkan performa model dalam hal memprediksi.

Anung Suprayitno, Kepala BMKG Stasiun Klimatologi Klas II Jawa Timur menyampaikan bahwa Jawa Timur dibagi kedalam 74 Zona Musim (ZOM) [13] yaitu daerah dengan pola rata-rata curah hujan menunjukkan perbedaan yang jelas antara musim hujan dan musim kemarau. Luas suatu ZOM tidak selalu sebanding dengan luas wilayah administrasi pemerintahan. Akibatnya, satu ZOM dapat meliputi beberapa kabupaten/kota, dan satu kabupaten/kota juga dapat terbagi ke dalam beberapa ZOM. Salah satu contohnya adalah ZOM 303 yang meliputi bagian utara Sidoarjo dan sebagian besar Kota Surabaya. Selain itu, topografi di Jawa Timur juga beragam, mulai dari pegunungan yang tinggi, dataran rendah, hingga pesisir [21] yang juga turut memengaruhi pola datangnya musim hujan dan kemarau. Variasi ini menyebabkan perbedaan dalam durasi dan intensitas musim di berbagai wilayah, sehingga memerlukan sistem prediksi yang lebih baik.

Penelitian ini menggunakan uji sampel yang dilakukan pada tiga Zona Musim (ZOM) di wilayah Jawa Timur, yaitu ZOM 303, ZOM 349, dan ZOM 311. Pemilihan zona-zona ini didasarkan pada representasi topografi dan karakteristik curah hujan yang beragam di Jawa Timur. Dengan pendekatan uji sampel ini, penelitian dapat mengidentifikasi variasi pola curah hujan dan periode musim di wilayah yang berbeda. Selain itu, penelitian ini juga menghasilkan sebuah *Graphical User Interface* (GUI) yang memungkinkan pengguna, termasuk masyarakat umum, untuk mengakses dan menginterpretasi hasil prediksi dengan mudah.

Penelitian ini diharapkan mampu untuk memberikan kontribusi langsung kepada masyarakat lokal yang bergantung pada informasi cuaca, terutama dalam sektor pertanian, perikanan, sumber daya air, dan infrastruktur. Namun, tantangan dalam memodelkan pola curah hujan yang kompleks di berbagai zona musim tetap menjadi perhatian utama. Oleh karena itu, penelitian ini difokuskan secara spesifik untuk menghasilkan prediksi periode musim yang lebih relevan dengan kondisi lokal, sehingga dapat memberikan manfaat praktis bagi pengambilan keputusan. Hasil penelitian ini juga diharapkan menjadi dasar bagi pengembangan model prediksi yang lebih canggih dan komprehensif, serta dapat diintegrasikan dengan sistem peringatan dini bencana guna meningkatkan kesiapsiagaan masyarakat terhadap perubahan iklim dan potensi bencana.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, dapat diidentifikasi masalah-masalah sebagai berikut.

- a. Bagaimana implementasi model *Stacked Long Short-Term Memory* (LSTM) dalam memprediksi curah hujan dasarian untuk menentukan periode musim di wilayah Jawa Timur?
- b. Bagaimana evaluasi kinerja model *Stacked LSTM* berdasarkan metrik *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Error* (MAE)?
- c. Bagaimana hasil prediksi curah hujan dasarian yang dihasilkan oleh model *Stacked LSTM*?
- d. Bagaimana penerapan hasil prediksi berbasis model *Stacked LSTM* ke dalam antarmuka grafis (GUI) interaktif menggunakan Streamlit?

1.3. Batasan Masalah

Berdasarkan rumusan masalah, maka terdapat batasan dalam melakukan penelitian diantaranya:

- a. Penelitian ini menggunakan data historis curah hujan, rata-rata suhu, dan kecepatan angin harian yang dikelompokkan per sepuluh hari atau dasarian dalam rentang tahun 2000 - 2024, sehingga prediksi yang dihasilkan merupakan prediksi dasarian tanpa detail prediksi dalam jam.
- b. Penelitian ini fokus pada analisis dan prediksi periode musim di beberapa wilayah Jawa Timur, yaitu Kepulauan Bawean Gresik, sebagian besar Kota Surabaya, bagian utara Sidoarjo, bagian timur laut Kabupaten Blitar, bagian tengah Kabupaten Malang, dan seluruh Kota Malang.
- c. Penelitian ini tidak mencakup analisis faktor-faktor lain yang mungkin mempengaruhi seperti vegetasi dan aktivitas manusia.
- d. Penelitian ini hanya melakukan perbandingan model *Stacked LSTM* dengan *Vanilla LSTM* dan *Bidirectional LSTM*.

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

- a. Untuk mengetahui implementasikan model *Stacked Long Short-Term Memory* (LSTM) dalam memprediksi curah hujan dasarian guna menentukan periode musim di wilayah Jawa Timur.
- b. Untuk mengetahui evaluasi kinerja model *Stacked LSTM* berdasarkan metrik *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Error* (MAE).
- c. Untuk mengetahui hasil prediksi curah hujan dasarian yang dihasilkan oleh model *Stacked LSTM*.
- d. Untuk mengetahui penerapan hasil prediksi model *Stacked LSTM* dalam bentuk antarmuka grafis (GUI) interaktif menggunakan Streamlit.

1.5. Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini, di antaranya:

- a. Dapat membantu petani menentukan waktu tanam dan panen yang optimal berdasarkan prediksi periode musim, sehingga dapat meningkatkan hasil produksi dan mengurangi risiko gagal panen.
- b. Dapat membantu menentukan waktu terbaik untuk proyek konstruksi, seperti pembangunan jalan, jembatan, atau gedung, dengan mempertimbangkan kondisi cuaca yang diprediksi.
- c. Untuk mendukung pemerintah dan masyarakat dalam merencanakan distribusi air yang efisien selama musim kemarau dan hujan.
- d. Untuk memberikan informasi awal terkait potensi bencana seperti banjir atau kekeringan, sehingga dapat dilakukan tindakan pencegahan.
- e. Dapat menyediakan referensi bagi penelitian serupa di wilayah lain, yang dapat membantu dalam pengembangan sistem peringatan dini terkait perubahan musim dan perencanaan.

Halaman ini sengaja dikosongkan