

BAB I

PENDAHULUAN

A. Latar Belakang

Indonesia merupakan negara dengan iklim tropis yang memiliki karakteristik curah hujan yang tinggi. Hal ini disebabkan oleh letak geografis yang berbeda diantara dua benua dan dua samudra serta dilalui garis khatulistiwa. Curah hujan sebagai salah satu unsur utama iklim memiliki peran penting dalam siklus air dan menjaga keseimbangan lingkungan. Variabilitas curah hujan yang tinggi di wilayah tropis dapat mengakibatkan perubahan dalam ketersediaan air serta perubahan debit sungai. Dalam kondisi tertentu, peningkatan intensitas dan akumulasi curah hujan dapat memicu terjadinya genangan hingga banjir, terutama ketika tanah tidak dapat menyerap air dengan baik dan sistem drainase tidak dapat menampung aliran air [1].

Kecamatan Serpong yang terletak di Kota Tangerang Selatan merupakan salah satu wilayah perkotaan yang mengalami perkembangan pesat dan memiliki tingkat kerentanan tinggi terhadap perubahan pola curah hujan. Kota Tangerang Selatan terletak di bagian timur Provinsi Banten, dan secara administratif terdiri dari tujuh kecamatan dan 54 kelurahan, salah satunya adalah Kecamatan Serpong [2]. Beberapa wilayah di Serpong diketahui rawan banjir, terutama di kawasan dengan drainase terbatas dan permukiman padat. Kecamatan Serpong termasuk wilayah dengan topografi datar dengan kemiringan tanah yang berkisar antara 0-3% serta elevansi di antara 0-25 meter di atas permukaan laut. Sebagian besar wilayahnya berada di Daerah Aliran Sungai (DAS) utama, yaitu Cisadane yang secara keseluruhan berkontribusi pada peningkatan risiko genangan dan banjir [3]. Kondisi ini mengakibatkan air hujan sulit teralirkan secara optimal, terutama pada kawasan dengan kepadatan permukiman yang tinggi dan kapasitas drainase yang terbatas.

Kerentanan tersebut tercermin dari kejadian banjir yang berulang di Kota Tangerang Selatan. Salah satu peristiwa yang signifikan terjadi pada 11 November 2024, ketika hujan dengan intensitas tinggi dalam waktu singkat memicu banjir di Kecamatan Serpong, terutama di Kelurahan Cilenggang dan Pondok Aren. Berdasarkan laporan Badan Penanggulangan Bencana Daerah (BPBD) Kota

Tangerang Selatan, curah hujan pada hari tersebut mencapai sekitar 150 mm/hari dan telah melebihi kapasitas sistem *drainase* perkotaan. Akibatnya, banjir dengan ketinggian 50–150 cm menggenangi kawasan permukiman, ruas jalan utama, serta fasilitas umum. Kejadian ini mengakibatkan terganggunya aktivitas masyarakat, mobilitas transportasi, serta kegiatan pendidikan [4]. Selain itu data BPBD Kota Tangerang Selatan menunjukkan peningkatan frekuensi kejadian banjir dari 54 kejadian pada tahun 2021 menjadi 203 kejadian pada tahun 2022, dengan jumlah warga yang terdampak mencapai 4.192 kepala keluarga [5]. Kondisi ini menunjukkan kerentanan Kecamatan Serpong terhadap curah hujan ekstrem. Mengingat banjir dapat terjadi dalam waktu singkat akibat hujan dengan intensitas tinggi, diperlukan model prediksi yang dapat memprediksi curah hujan dalam jangka waktu dekat, yaitu 30 menit kedepan, sehingga dapat digunakan sebagai dasar peringatan dini dan mitigasi risiko bencana di Kecamatan Serpong.

Meskipun penelitian mengenai prediksi curah hujan telah banyak dilakukan, penerapan model prediktif pada karakteristik geografis dan klimatologi Kecamatan Serpong masih terbatas. Kompleksitas hubungan urbanisasi, pola aliran air, dan kondisi *drainase* menyebabkan pola curah hujan bersifat *non-linier* dan sulit dimodelkan dengan menggunakan metode statistika konvensional. Metode statistika umumnya lebih sesuai untuk data berpola *linier* dan cenderung sensitif terhadap *outlier* [6]. Oleh karena itu, algoritma *machine learning* memiliki potensi lebih besar dalam menghasilkan prediksi curah hujan yang akurat.

Metode *ensemble learning* seperti *Random Forest* dan *XGBoost* dipilih karena memiliki keunggulan dalam menangani data berdimensi tinggi, mengidentifikasi hubungan *non-linear*, serta mengevaluasi kontribusi relatif setiap variabel prediktor [7]. Berdasarkan penelitian oleh Sangaji dan Sutabri [8], model gabungan *Random Forest* dan *XGBoost* berbasis *weighted average* menunjukkan performa lebih baik dibandingkan dengan model tunggal dalam memprediksi curah hujan di Sumatera Selatan. Penelitian tersebut menggunakan data *historis* curah hujan Sumatera Selatan tahun 2015-2023. Model gabungan dengan bobot *XGBoost* sebesar 0,6 dan *Random Forest* sebesar 0,4 menghasilkan nilai error terkecil yaitu MSE 14.125,78, RMSE 118,85, dan MAPE 11,87%. Temuan ini mengindikasikan bahwa penggabungan beberapa model dapat meningkatkan akurasi prediksi.

Berdasarkan indikasi tersebut, penelitian lebih lanjut diperlukan untuk mengeksplorasi teknik penggabungan model yang lebih optimal guna meningkatkan akurasi. Metode *Stacking* menawarkan pendekatan lebih adaptif dibandingkan *weighted average* karena mengkombinasikan prediksi dari *base learner* melalui *meta learner* yang dilatih untuk menghasilkan prediksi akhir dengan kinerja lebih baik [7]. Penelitian Nalatissifa [9], mengimplementasikan *Stacking ensemble learning* untuk klasifikasi produktivitas tenaga kerja di industri garmen dari Kaggle. Studi ini menggabungkan tiga *base learner* yaitu *XGBoost*, *Random Forest*, dan *LightGBM*. Hasilnya menunjukkan performa klasifikasi sangat tinggi dengan akurasi 99,58%, presisi 100%, dan recall 99,29%. Hal ini menunjukkan metode *stacking* efektif mengintegrasikan hasil prediksi *base learner* untuk meningkatkan kinerja model.

Oleh karena itu, untuk mengatasi kerentanan banjir di Kecamatan Serpong yang dipicu oleh curah hujan ekstrem, serta memanfaatkan peluang peningkatan akurasi dari teknik *ensemble*, penelitian ini mengusulkan untuk mengembangkan model prediksi curah hujan dengan teknik *Stacking* yang mengombinasikan *Random Forest* dan *XGBoost*. Diharapkan model ini dapat memberikan prediksi curah hujan 30 menit kedepan yang lebih akurat, sehingga dapat digunakan sebagai sistem peringatan dini dan mendukung perencanaan tata kota yang lebih efektif di Kecamatan Serpong, Kota Tangerang Selatan.

B. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan rumusan masalah dalam penelitian adalah “Bagaimana merancang dan mengimplementasikan metode *stacking ensemble* dengan *Random Forest* dan *XGBoost* sebagai *base-learner*, serta menerapkan *Random Forest* dan *XGBoost* sebagai *meta-learner* menggunakan hasil prediksi *base-learner* sebagai input dalam memprediksi curah hujan di Kecamatan Serpong, Kota Tangerang Selatan?”.

C. Batasan Penelitian

Batasan masalah yang ditentukan dalam penelitian ini mencakup beberapa hal sebagai berikut:

1. Dataset yang digunakan merupakan Data Curah Hujan dari Stasiun AQMS Kota Tangerang Selatan tahun 2024 dalam bentuk data harian dengan rentang data per 30 menit. Dataset terdiri dari variabel kecepatan angin, arah angin, kelembaban, suhu, tekanan udara, radiasi matahari (Sol.Rad) dan curah hujan.
2. Optimasi *hyperparameter* menggunakan *Grid Search CV* hanya pada beberapa *hyperparameter* tertentu, yaitu jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), minimum sampel untuk membagi node (*min_samples_split*), dan minimum sampel pada node daun (*min_samples_leaf*) untuk *Random Forest*, serta jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), dan tingkat pembelajaran model (*learning_rate*) untuk *XGBoost*.

D. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk merancang dan mengimplementasikan model prediksi berbasis *stacking ensemble* dengan *Random Forest* dan *XGBoost* sebagai *base learner* dan *meta-learner* guna meningkatkan performa prediksi curah hujan di Kecamatan Serpong, Kota Tangerang Selatan.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat dan dampak positif dari penerapan model prediksi curah hujan kecamatan Serpong, Kota Tangerang Selatan dengan menggabungkan algoritma *Random Forest* dan *XGBoost* menggunakan teknik *stacking*. Secara umum, manfaat penelitian dibagi menjadi tiga aspek, yaitu:

1. Manfaat akademik

Penelitian ini diharapkan dapat memperkaya kajian ilmiah di bidang klimatologi dan penerapan *machine learning*, serta menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan prediksi curah hujan.

2. Manfaat bagi pemerintah

Hasil penelitian ini diharapkan dapat menyediakan informasi pendukung bagi pemerintah daerah dalam menyusun kebijakan mitigasi bencana banjir di kecamatan Serpong dan mendukung perencanaan tata kota yang lebih baik.

3. Manfaat bagi masyarakat

Penelitian ini diharapkan dapat meningkatkan tingkat kesiapan masyarakat dalam menghadapi potensi banjir akibat curah hujan tinggi.

