

BAB I

PENDAHULUAN

1.1 Latar Belakang

Proses kondensasi uap air di atmosfer menyebabkan terjadinya hujan. Hujan merupakan salah satu fenomena cuaca yang dipengaruhi oleh berbagai perubahan pada sejumlah variabel cuaca di atmosfer. Salah satu informasi penting yang berkaitan dengan fenomena hujan adalah intensitas hujan yang turun yang dapat disebut sebagai curah hujan (Gede et al., 2022). Curah hujan di Indonesia memiliki karakteristik unik yang dipengaruhi oleh faktor ketinggian daerah, iklim dan musim, yang mengakibatkan curah hujan di tiap daerah memiliki tingkat curah hujan yang berbeda (Ihsan et al., 2021).

Salah satu provinsi di Indonesia yang memiliki curah hujan relatif cukup tinggi adalah provinsi Sumatera Utara. Hal ini ditunjukkan pada peristiwa yang terjadi pada tahun 2020, intensitas curah hujan mencapai 4.380 mm yang tergolong tinggi dibandingkan curah hujan normal tahunan, yaitu 1000-3000 per tahun (BPS, 2020). Curah hujan yang tinggi ini menyebabkan Kota Medan yang merupakan Ibu kota provinsi Sumatra utara masih rentan terhadap banjir (Pasaribu, 2021).

Menurut BNPB, banjir terparah di Kota Medan terjadi pada tahun 2022. Data dari Stasiun Klimatologi Deli Serdang menunjukkan bahwa banjir tersebut terjadi sejak 27 Februari 2022, disebabkan oleh hujan deras yang mengakibatkan meluapnya sungai-sungai di Kota Medan. Berdasarkan informasi yang didapatkan melalui BMKG di wilayah Sumatera Utara, bulan Februari seharusnya sudah memasuki puncak kemarau pertama. Namun, realitasnya banjir tetap terjadi di Kota Medan (Paramita et al., 2023). Ketidaksesuaian ini menandakan pentingnya peramalan curah hujan yang akurat sebagai bagian dari upaya mitigasi risiko bencana banjir.

Agar peramalan curah hujan dapat dilakukan secara akurat, perlu mempertimbangkan berbagai variabel meteorologis yang saling berinteraksi dan

memengaruhi pembentukan hujan. Di antaranya adalah temperatur, kelembapan dan lama penyinaran matahari. Temperatur udara dipengaruhi oleh lamanya penyinaran matahari. Semakin lama matahari menyinari bumi maka semakin panas pula permukaan bumi (Deqita, 2022). Kelembapan udara bergantung pada penguapan air dari laut, danau, sungai, dan tanah. Suhu dan kelembapan juga berkaitan dan memiliki pengaruh terhadap curah hujan, jika suhu suatu daerah tinggi maka kelembapannya rendah begitu pula sebaliknya (Bili & Gowa, 2024). Penelitian Musfiroh, et al (2023) menunjukkan bahwa tekanan, evaporasi, temperatur maksimal, kelembapan rata-rata, dan lama penyinaran matahari merupakan variabel-variabel penting yang berpengaruh terhadap curah hujan. Oleh karena itu, pendekatan peramalan yang mampu menangkap kompleksitas hubungan antar variabel ini sangat dibutuhkan.

Keakuratan peramalan bergantung pada dua hal penting, yaitu pemilihan metode yang tepat dan ketersediaan data curah hujan yang lengkap. Pentingnya melakukan prediksi telah mendorong para peneliti untuk menciptakan teknik peramalan yang lebih tepat (Insani et al., 2020). Penggunaan teknik yang berbeda akan menghasilkan nilai akurasi yang berbeda pula. Oleh karena itu, penting untuk memilih algoritma yang sesuai dan memodelkannya sesuai dengan kebutuhan (Mahajan et al., 2022).

Seiring perkembangan teknologi, para peneliti menggunakan teknik *data mining*, analisis *big data*, dan berbagai algoritma *machine learning* untuk meningkatkan keakuratan akurasi dalam meramalkan curah hujan secara harian, bulanan, dan tahunan. Berdasarkan hasil penelitian yang telah dilakukan oleh para peneliti, algoritma *machine learning* memiliki tingkat efektivitas yang lebih tinggi dibandingkan teknik data mining konvensional untuk meramalkan curah hujan dan cuaca (Liyew et al., 2021). Keunggulan *machine learning* terletak pada kemampuannya dalam menangkap pola non-linear atau tidak teratur pada data (Poornima et al., 2023). Salah satu algoritma machine learning yang digunakan untuk memprediksi ataupun untuk regresi adalah XGBoost (*eXtreme Gradient Boosting*).

Algoritma XGBoost merupakan algoritma modern dan populer berdasarkan kerangka *Gradient Boosting Machine* (GBM). Algoritma ini mampu bekerja

dengan baik pada berbagai tugas, termasuk regresi, klasifikasi, dan pemeringkatan. (Anwar et al., 2021). Dengan memanfaatkan *ensemble learning*, XGBoost mengembangkan pohon keputusan dan memperbarui model untuk memperbaiki ketidakakuratan yang ditemukan pada model sebelumnya. XGBoost mampu menangani data berskala besar dengan efisien dan dapat memberikan hasil akurasi yang lebih tinggi karena menggunakan *decision tree* untuk mengidentifikasi variabel penting dan membangun model. XGBoost menjadi metode yang sering digunakan karena keakuratan, efisiensi, dan kemampuan dalam beradaptasinya (Tarwidi et al., 2023). Keunggulan algoritma ini adalah kecepatan dalam memproses dan kemampuan melakukan perhitungan di luar inti. Namun algoritma ini memiliki kelemahan yaitu kemungkinan *overfitting* yang dapat diatasi dengan bereksperimen menggunakan parameter pemodelan (Lisanthoni et al., 2023).

Penelitian terdahulu menyatakan bahwa XGBoost efektif dalam meramalkan curah hujan harian di Kota Semarang. Penelitian yang dilakukan oleh Anwar et al. (2021) memiliki hasil yang cukup baik dengan nilai RMSE sebesar 2,7 mm dan MAE sebesar 8,8 mm. Dalam penelitian ini menggunakan *feature importance* untuk mengidentifikasi variabel yang berpengaruh. Maka, didapatkanlah dua variabel yang sangat mempengaruhi peramalan curah hujan, yaitu kelembapan rata-rata dan temperatur minimum.

Liyew et al. (2021) dalam penelitiannya membahas teknik *machine learning* untuk memprediksi curah hujan harian menggunakan tiga algoritma, yaitu *Multivariate linear regression* (MLR), *Random Forest* (RF), dan XGBoost. Dalam penelitian ini, metode *Pearson Correlation* digunakan untuk memilih fitur-fitur yang memiliki korelasi kuat dengan variabel curah hujan. Hasil penelitian menunjukkan bahwa XGBoost mampu memberikan performa prediksi yang lebih baik dibandingkan MLR dan *Random Forest*.

Penelitian yang dilakukan oleh Muslim Karo Karo (2020) membahas mengenai penerapan metode XGBoost dan *feature importance* dalam klasifikasi kebakaran hutan dan lahan. Hasil penelitian menunjukkan bahwa penerapan *feature importance* mampu meningkatkan akurasi dengan mengidentifikasi enam hingga tujuh variabel yang memiliki pengaruh signifikan dalam menentukan keberadaan titik panas. XGBoost terbukti mampu memberikan performa yang baik, dengan

tingkat akurasi sebesar 89,52%. Hal ini menunjukkan bahwa penggunaan *feature importance* efektif dalam meningkatkan akurasi klasifikasi kebakaran hutan dan lahan.

Berdasarkan paparan diatas, penelitian ini akan melakukan peramalan intensitas curah hujan menggunakan metode XGBoost dan *feature importance*. Penggunaan *feature importance* bertujuan untuk mengidentifikasi variabel-variabel yang memiliki pengaruh signifikan terhadap curah hujan. Dengan pendekatan ini, diharapkan peramalan curah hujan yang dihasilkan mendapatkan hasil akurasi yang lebih baik, sehingga pemerintah dan masyarakat dapat mengambil langkah-langkah yang preventif dan mitigasi dalam menghadapi potensi dampak curah hujan yang tinggi di Kota Medan.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dipaparkan, beberapa permasalahan utama yang dapat diidentifikasi dalam penelitian ini, sebagai berikut.

1. Kota Medan sering mengalami banjir akibat tingginya curah hujan, terutama pada periode tertentu. Oleh karena itu, peramalan curah hujan yang akurat diperlukan untuk mendukung upaya mitigasi banjir.
2. Penelitian sebelumnya umumnya hanya memanfaatkan data historis curah hujan sebagai variabel tunggal, tanpa mempertimbangkan faktor-faktor meteorologis lain seperti temperatur, kelembapan udara, dan lama penyinaran matahari yang juga berpengaruh terhadap curah hujan.
3. Model peramalan curah hujan di Kota Medan pada penelitian terdahulu menunjukkan tingkat akurasi yang belum optimal, sehingga hasil prediksinya kurang dapat diandalkan dalam mendukung upaya mitigasi banjir.
4. Masih terbatasnya penggunaan metode yang mampu menangani data berskala besar dan mengenali pola non-linear yang kompleks akibat interaksi antar variabel meteorologis menjadi tantangan dalam membangun model peramalan yang akurat.

5. Penggunaan teknik *feature importance* untuk mengidentifikasi variabel-variabel yang paling berpengaruh terhadap curah hujan masih jarang diterapkan dalam penelitian serupa.

1.3 Rumusan Masalah

Berdasarkan uraian latar belakang, rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana meramalkan intensitas curah hujan di Kota Medan berdasarkan faktor-faktor yang mempengaruhi curah hujan?
2. Bagaimana mengimplementasikan metode XGBoost menggunakan pemrograman *Python* dalam meramalkan curah hujan di Kota Medan?
3. Bagaimana mengevaluasi kinerja metode XGBoost dalam meramalkan curah hujan di Kota Medan?
4. Bagaimana meningkatkan akurasi metode XGBoost dalam peramalan curah hujan di Kota Medan?

1.4 Batasan Masalah

Adapun batasan masalah dalam penelitian ini, sebagai berikut.

1. Data yang digunakan dalam penelitian ini merupakan data *time series* curah hujan harian dari tahun 2013 hingga tahun 2022 di Kota Medan.
2. Data yang digunakan dalam penelitian ini berasal dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Wilayah I Medan.
3. Variabel yang digunakan dalam penelitian ini, meliputi: curah hujan (mm), lama penyinaran matahari (jam), temperatur minimum (°C), temperatur maksimum (°C), temperatur rata-rata (°C), dan kelembapan rata-rata (%).
4. Evaluasi performa metode XGBoost dalam peramalan intensitas curah hujan menggunakan MAE dan RMSE.

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, maka tujuan dari penelitian ini adalah sebagai berikut.

1. Meramalkan intensitas curah hujan di Kota Medan berdasarkan faktor-faktor yang mempengaruhi curah hujan menggunakan metode XGBoost.
2. Mengimplementasikan metode XGBoost menggunakan pemrograman *python* sesuai dengan alur kerja metode XGBoost.
3. Mengevaluasi kinerja metode *XGBoost* dalam peramalan curah hujan di Kota Medan menggunakan MAE dan RMSE.
4. Menganalisis pengaruh masing-masing variabel terhadap hasil prediksi menggunakan *feature importance* untuk meningkatkan akurasi metode XGBoost dalam peramalan curah hujan di Kota Medan.

1.6 Manfaat Penelitian

Adapun manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut.

1. Bagi peneliti, penelitian ini untuk menambah wawasan mengenai penerapan metode XGBoost dalam peramalan curah hujan, khususnya di wilayah Kota Medan.
2. Bagi pembaca, hasil penelitian ini dapat menjadi referensi tambahan dalam mengembangkan penelitian yang menggunakan metode XGBoost untuk peramalan data cuaca atau data *time series* lainnya..
3. Bagi instansi terkait, khususnya Badan Meteorologi Klimatologi dan Geofisika (BMKG) Sumatera Utara, penelitian ini diharapkan dapat memberikan masukan terkait penerapan metode XGBoost sebagai salah satu pendekatan yang dapat digunakan dalam meningkatkan akurasi peramalan curah hujan.