

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan kecerdasan buatan telah membuka peluang luas dalam pemanfaatan data historis untuk peramalan di berbagai bidang [1]. Salah satu tantangan fundamental dalam domain keinformatikaan adalah bagaimana mengadaptasi algoritma *machine learning* regresi umum, seperti *Random Forest Regression* dan *Support Vector Regression*, agar mampu digunakan pada data runtun waktu (*time series*) [2]. Kedua algoritma tersebut pada dasarnya tidak dirancang untuk mengenali pola urutan waktu secara inheren, sehingga tidak dapat langsung menangkap informasi temporal (*sekuensial*) yang terkandung dalam data. Hal ini menimbulkan kebutuhan akan pendekatan khusus yang memungkinkan algoritma *non-time series* tetap dapat digunakan untuk tugas peramalan (*forecasting*).

Pendekatan rekayasa fitur (*feature engineering*) memungkinkan transformasi data deret waktu menjadi data *supervised learning* melalui penciptaan fitur tambahan, seperti *lag* (nilai masa lalu), representasi musiman (siklus bulanan) [3], dan statistik historis (*mean*, *median*, atau deviasi standar per periode) [4]. Dengan cara ini, algoritma regresi yang awalnya “buta waktu” dapat “diajarkan” untuk mengenali pola temporal dan tren musiman, sehingga dapat digunakan untuk meramalkan nilai masa depan. Pendekatan ini memiliki keunggulan dalam hal fleksibilitas, kemampuan menangkap pola non-linear [5] serta kemudahan implementasi pada berbagai jenis data tanpa memerlukan infrastruktur komputasi yang kompleks seperti model numerik berbasis fisika atmosfer [6].

Di sisi lain, fenomena ketidakstabilan curah hujan di Indonesia memberikan dampak signifikan pada berbagai sektor vital seperti pertanian, sumber daya air, transportasi, dan ketahanan pangan [7]. Ketidakpastian ini menyulitkan pengambilan keputusan dan meningkatkan risiko kerugian bagi para pelaku di sektor terkait. Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) memproyeksikan sekitar 67% wilayah Indonesia akan mengalami curah hujan di atas 2.500 mm per tahun pada periode 2024–2025, bahkan beberapa wilayah seperti Sulawesi, Papua, dan

Kalimantan diperkirakan melebihi 5.000 mm per tahun [8]. Kondisi ini menekankan kebutuhan akan sistem Peramalan curah hujan yang akurat dan dapat diakses secara lebih luas.

Di Indonesia, Peramalan curah hujan umumnya dilakukan oleh Badan BMKG dengan memanfaatkan teknologi seperti satelit, radar, dan model numerik berbasis fisika atmosfer [9]. Metode ini dinilai efektif karena mampu menangkap dinamika atmosfer secara luas dan akurat melalui dukungan perangkat keras seperti satelit dan radar cuaca [10]. Namun, penggunaan teknologi ini memerlukan infrastruktur perangkat keras dan sistem komputasi yang kompleks serta keahlian teknis khusus, sehingga belum tentu dapat diadopsi secara luas oleh pihak di luar lembaga resmi [5].

Sebagai alternatif, penggunaan *machine learning* berbasis data historis menawarkan solusi yang lebih fleksibel dan terjangkau. Namun, masih terdapat celah penelitian (*research gap*), khususnya dalam konteks Indonesia, terkait evaluasi komparatif performa algoritma RF dan SVR pada data curah hujan jangka panjang di wilayah tertentu. Banyak penelitian sebelumnya hanya berfokus pada salah satu algoritma, menggunakan data multiregional, atau periode pengamatan yang relatif pendek.

Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan mengembangkan dan membandingkan model peramalan curah hujan bulanan berbasis RFR dan SVR secara komprehensif di Kota Bandung. Dengan fokus pada pendekatan univariat dan penerapan rekayasa fitur, penelitian ini diharapkan dapat menghasilkan model yang akurat, efisien, serta memberikan kontribusi metodologis dalam mendukung adaptasi algoritma regresi umum untuk tugas peramalan data deret waktu (*time series*).

1.2 Perumusan Masalah Penelitian

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, yang dapat dirumuskan masalah dalam penelitian ini adalah:

1. Bagaimana membangun model Peramalan curah hujan menggunakan algoritma *Random Forest Regression* dan *Support Vector Regression* (SVR)?

2. Bagaimana perbandingan performa Algoritma *Random Forest Regression* dan *Support Vector Regression* (SVR) yang memiliki performa lebih baik dalam melakukan Peramalan curah hujan?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk:

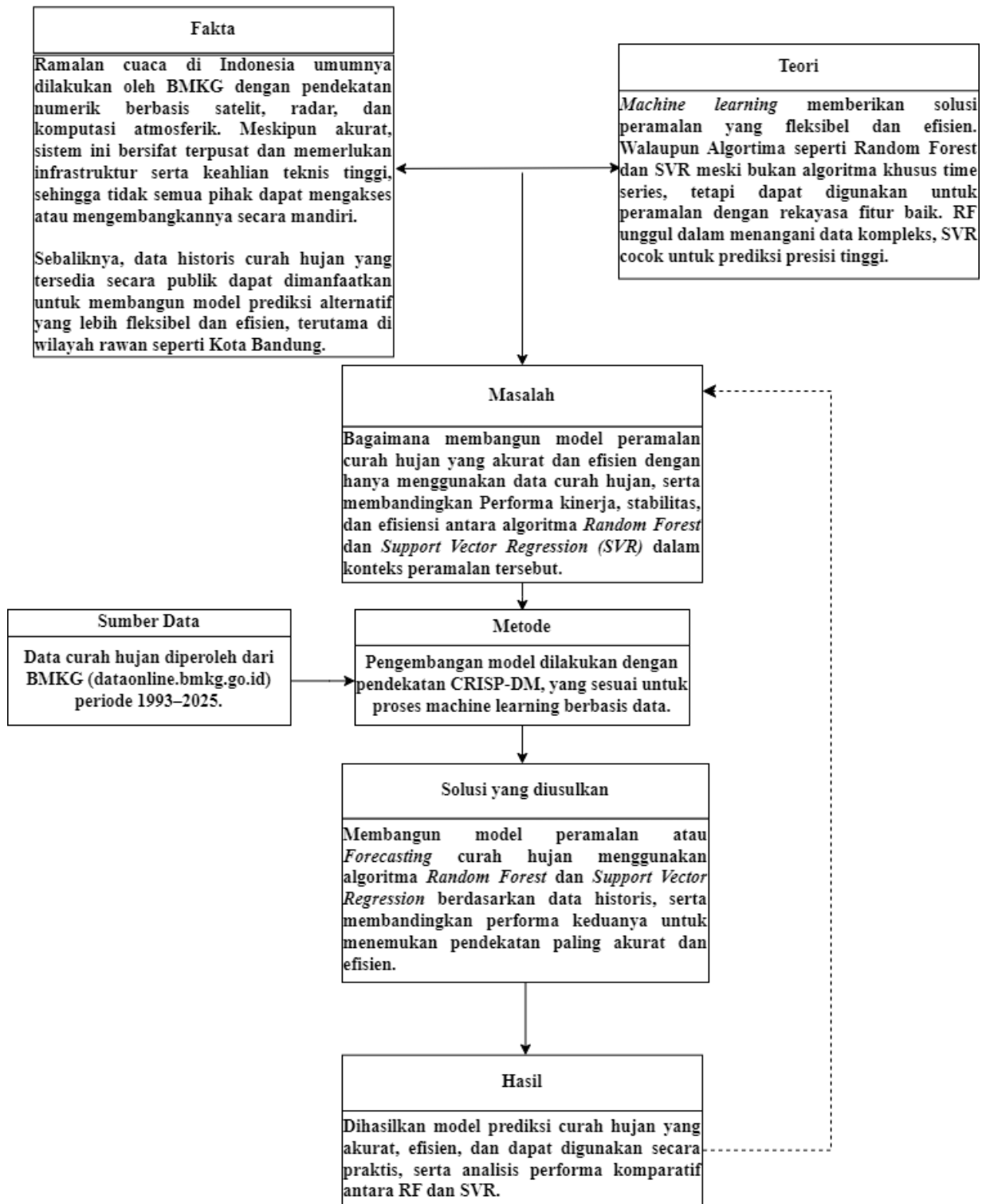
1. Membangun model peramalan curah hujan menggunakan algoritma *Random Forest Regression* dan *Support Vector Regression* (SVR)
2. Membandingkan performa algoritma *Random Forest Regression* dan *Support Vector Regression* (SVR) dalam meramalkan curah hujan.

1.4 Batasan Masalah Penelitian

Agar penelitian ini tetap berada dalam ruang lingkup yang telah ditentukan, maka ditetapkan beberapa batasan masalah sebagai berikut:

1. Variabel yang digunakan dalam proses pelatihan dan Peramalan hanya berupa data curah hujan dalam satuan mm (milimeter)
2. Penelitian ini menggunakan data sekunder berupa data historis curah hujan bulanan dari situs [dataonline.bmkg](http://dataonline.bmkg.go.id) Stasiun Geofisika BMKG Bandung. Cakupan data meliputi wilayah Kota Bandung untuk periode waktu dari Januari 1993 hingga Maret 2025.
3. Model memberikan hasil Peramalan berupa estimasi nilai curah hujan dalam satuan milimeter (mm) beserta grafik visualisasi curah hujan.
4. Model hanya bisa melakukan peramalan untuk satu tahun kedepan dari titik data terakhir dalam dataset historis yang digunakan.

1.5 Kerangka Pemikiran Penelitian



Gambar 1. 1 Kerangka Pemikiran Penelitian

Gambar 1.1 menunjukkan alur penelitian yang dimulai dari fakta bahwa sistem Peramalan cuaca di Indonesia, seperti yang dikembangkan oleh BMKG, menggunakan pendekatan numerik berskala besar dan dikelola secara terpusat. Meskipun akurat, sistem tersebut memerlukan infrastruktur dan keahlian teknis tertentu, sehingga tidak semua pihak dapat mengakses atau mengembangkannya secara mandiri. Sementara itu, data historis curah hujan yang tersedia secara publik membuka peluang untuk pengembangan model Peramalan alternatif yang lebih fleksibel. Berdasarkan teori bahwa *machine learning* mampu mengolah pola historis secara efisien, penelitian ini menggunakan algoritma *Random Forest Regression* (RFR) dan *Support Vector Regression* (SVR) untuk membangun model Peramalan curah hujan. Proses pengembangan mengikuti metodologi CRISP-DM, dengan memanfaatkan data curah hujan wilayah Kota Bandung periode 1993–2025. Penelitian ini menghasilkan model Peramalan yang akurat dan efisien, serta menyajikan analisis perbandingan performa antara algoritma RFR dan SVR dari aspek akurasi, stabilitas, dan efisiensi.

1.6 Sistematis Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang proses penulisan yang dilakukan untuk menyelesaikan penelitian ini. Sistematika penulisan penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penelitian, perumusan masalah, tujuan penelitian, serta batasan yang ditetapkan agar penelitian tetap fokus. Selain itu, bab ini juga menjelaskan kerangka pemikiran yang menjadi dasar penelitian, metodologi yang digunakan, serta gambaran sistematika penulisan sebagai panduan dalam membaca penelitian ini.

BAB II STUDI PUSTAKA

Bab ini membahas teori-teori dan konsep dasar yang relevan dengan topik penelitian. Kajian pustaka ini bertujuan untuk memberikan pemahaman mendalam serta mendukung analisis dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan langkah-langkah metodologi yang digunakan dalam penelitian, mulai dari tahap awal hingga akhir. Pembahasan mencakup analisis produk serta perencanaan eksekusi aplikasi agar penelitian berjalan secara sistematis dan terarah.

BAB IV HASIL DAN PEMBAHASAN

Bab ini terdiri dari dua hal utama, pertama pemaparan tentang temuan atau hasil penelitian berdasarkan tahapan penelitian yang dilakukan. Pemaparan hasil penelitian disesuaikan dengan urutan masalah penelitian. Kedua pembahasan hasil atau temuan penelitian untuk menjawab rumusan penelitian.

BAB V PENUTUP

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan, serta kritik dan saran yang dapat menjadi masukan bagi penelitian selanjutnya. Kesimpulan yang disajikan merupakan intisari dari keseluruhan penelitian dan dapat dijadikan dasar pengembangan lebih lanjut.

