

BAB I

PENDAHULUAN

1.1. Latar Belakang

Curah hujan merupakan salah satu variabel klimatologis yang memiliki peran strategis dalam menjaga keseimbangan ekosistem serta mendukung sektor pertanian, pengelolaan sumber daya air, dan mitigasi bencana hidrometeorologi [1]. Informasi mengenai curah hujan sangat penting karena berkaitan langsung dengan berbagai aktivitas manusia yang bergantung pada kondisi cuaca dan iklim [2]. Selain itu, data curah hujan juga menjadi salah satu indikator yang digunakan dalam perencanaan pembangunan dan pengelolaan lingkungan secara berkelanjutan [3].

Sebagai negara beriklim tropis, Indonesia memiliki karakteristik curah hujan yang dinamis dan bervariasi antarwilayah maupun antarperiode waktu [4]. Variabilitas tersebut dipengaruhi oleh berbagai faktor atmosfer dan oseanografi, seperti sistem monsun, El Niño Southern Oscillation (ENSO), dan Indian Ocean Dipole (IOD) [5]. Kondisi geografis Indonesia yang terdiri atas ribuan pulau dengan topografi yang beragam turut memengaruhi distribusi curah hujan pada setiap wilayah [6]. Oleh karena itu, analisis terhadap pola curah hujan menjadi penting untuk memahami karakteristik iklim di Indonesia [7].

Perubahan pola curah hujan dapat memberikan dampak yang signifikan terhadap berbagai sektor kehidupan [8]. Curah hujan yang tinggi berpotensi menyebabkan banjir dan tanah longsor di berbagai daerah [9]. Sebaliknya, penurunan curah hujan dalam periode tertentu dapat memicu kekeringan yang berdampak pada sektor pertanian dan ketersediaan air bersih [10]. Ketidakpastian curah hujan juga dapat menyebabkan kerugian ekonomi serta memengaruhi ketahanan pangan masyarakat [11]. Oleh karena itu, diperlukan pemahaman yang lebih komprehensif terhadap pola curah hujan sebagai dasar dalam pengambilan keputusan yang berbasis data [12].

Perkembangan teknologi pengamatan cuaca dan iklim telah menghasilkan peningkatan volume data klimatologi yang sangat besar [13]. Data yang dikumpulkan secara rutin oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) terus bertambah setiap waktu dan tersimpan dalam berbagai format data [14]. Karakteristik data yang semakin kompleks menimbulkan tantangan dalam proses penyimpanan, pengelolaan, dan analisis data [15]. Pengolahan data menggunakan metode konvensional sering kali memerlukan waktu yang lebih lama ketika dihadapkan pada data dalam jumlah besar [16]. Kondisi tersebut mendorong perlunya pemanfaatan teknologi yang mampu mengelola data secara lebih efektif dan efisien [17].

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah teknologi Big Data [18]. Teknologi Big Data memungkinkan proses pengolahan data berskala besar dilakukan dengan lebih cepat dibandingkan pendekatan tradisional [19].

Dalam ekosistem Big Data, Apache Spark menjadi salah satu framework yang banyak digunakan karena mendukung pemrosesan data secara terdistribusi [20]. Apache Spark juga memiliki kemampuan in-memory computing yang dapat meningkatkan performa pengolahan data [21]. Selain itu, PySpark menyediakan berbagai fungsi yang mendukung proses transformasi dan analisis data menggunakan bahasa pemrograman Python [22].

Beberapa penelitian sebelumnya telah menunjukkan bahwa teknologi Big Data mampu meningkatkan efektivitas analisis data meteorologi dan klimatologi [23]. Penelitian lain juga menunjukkan bahwa Apache Spark dapat dimanfaatkan untuk mengolah dataset berukuran besar secara efisien [24]. Namun, penelitian yang secara khusus membahas analisis deskriptif data curah hujan Indonesia menggunakan Apache Spark dengan memanfaatkan data resmi BMKG masih relatif terbatas [25]. Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk melakukan analisis data curah hujan menggunakan pendekatan Big Data berbasis Apache Spark guna memperoleh informasi mengenai karakteristik curah hujan di Indonesia serta mengevaluasi pemanfaatan Apache Spark dalam pengolahan data klimatologi berskala besar.

1.2. Identifikasi Masalah

Dalam konteks analisis curah hujan menggunakan teknologi Big Data, terdapat beberapa isu krusial yang membutuhkan perhatian dan pemecahan. Masalah-masalah yang teridentifikasi yaitu :

1. Minimnya Pemanfaatan Data Meteorologi

Meskipun Indonesia memiliki data cuaca yang melimpah dari BMKG, pemanfaatannya dalam penelitian dan perumusan kebijakan masih belum optimal.

2. Kompleksitas dan Volume Data

Data curah hujan berskala besar dan memiliki banyak variabel, sehingga analisis manual menjadi tidak efisien dan memerlukan teknologi Big Data.

3. Keterbatasan Pemahaman Pola Curah Hujan

Variabilitas curah hujan menimbulkan ketidakpastian dalam perencanaan pertanian dan pengelolaan irigasi.

4. Kendala Adopsi Teknologi

Masih terdapat keterbatasan dalam penerapan teknologi Big Data di instansi pemerintah dan sektor pertanian.

5. Terbatasnya Studi Berbasis Big Data

Penelitian yang mengolah data iklim menggunakan pendekatan Big Data di Indonesia masih relatif sedikit.

6. Dampak Sosial dan Ekonomi

Ketidakpastian curah hujan dapat berdampak pada kerugian ekonomi, ketahanan pangan, dan kerentanan sosial masyarakat agraris.

1.3. Batasan Masalah

Agar penelitian ini lebih terfokus dan dapat memberikan hasil yang optimal, beberapa batasan masalah telah ditetapkan sebagai berikut:

1.Fokus Data Iklim

Penelitian hanya menganalisis data curah hujan, suhu, dan kelembapan dari dataset resmi BMKG. Data lain seperti ekonomi atau demografi tidak disertakan.

2.Sumber dan Format Data

Data yang digunakan terbatas pada dataset BMKG berformat CSV dan time series dengan rentang waktu serta lokasi tertentu. Tidak menggunakan sumber data lain.

3.Pendekatan Deskriptif

Analisis bersifat deskriptif untuk mengidentifikasi pola historis tanpa membangun model prediktif.

4.Tools yang Digunakan

Penelitian ini menggunakan Apache Spark (PySpark) dan Python untuk pengolahan serta analisis data. Platform lain seperti R atau MATLAB tidak digunakan agar metodologi tetap fokus dan konsisten

5.Batasan Wilayah

Penelitian hanya mencakup wilayah tertentu di Indonesia sehingga hasilnya tidak dapat digeneralisasi secara nasional tanpa studi lanjutan.

6.Batasan Waktu

Data terbatas pada periode tertentu dan mungkin belum merepresentasikan perubahan iklim jangka panjang.

7.Tidak Mempertimbangkan Faktor Eksternal

Faktor seperti kebijakan pemerintah, perubahan iklim global, atau aspek sosial ekonomi tidak dianalisis.

8.Tidak Ada Perbandingan Teknologi

Penelitian tidak membandingkan Spark dengan platform Big Data lainnya; pemilihan Spark didasarkan pada kebutuhan pengolahan data skala besar.

1.4. Rumusan Masalah

Berdasarkan latar belakang, identifikasi masalah, dan batasan yang telah ditetapkan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Apa saja pola curah hujan yang dapat diidentifikasi dari data yang diperoleh dari BMKG?
2. Bagaimana efektivitas penggunaan Apache Spark dalam mengolah dan menganalisis data curah hujan berukuran besar?
3. Apa dampak dari pemahaman pola curah hujan terhadap kebijakan pertanian di Indonesia?
4. Bagaimana hubungan antara curah hujan dengan variabel iklim lainnya, seperti suhu dan kelembapan?
5. Apa tantangan yang dihadapi dalam pengumpulan dan analisis data curah hujan dalam penelitian ini?

1.5 Manfaat Penelitian

1. Bagi Peneliti

Penelitian ini memberikan pemahaman yang mendalam tentang aplikasi teknologi Big Data, khususnya melalui penggunaan Apache Spark dalam analisis data iklim.

2. Bagi Pemerintah

Hasil penelitian ini akan menyediakan informasi yang berbasis data untuk pengambilan keputusan strategis dalam kebijakan pertanian, serta rencana mitigasi bencana.

3. Bagi Sektor Pertanian

Analisis pola curah hujan yang terperinci akan membantu petani dalam mengoptimalkan waktu penanaman dan strategi pengelolaan lahan.

4. Bagi Masyarakat

Penelitian ini diharapkan dapat meningkatkan kesadaran masyarakat tentang pentingnya memahami pola curah hujan dan dampaknya terhadap kehidupan sehari-hari, seperti pengelolaan sumber daya air dan adaptasi terhadap perubahan iklim.

5. Bagi Akademisi

Penelitian ini akan menjadi referensi yang berharga bagi akademisi yang mempelajari iklim dan cuaca, serta memberikan wawasan tentang penggabungan teknologi modern dengan metodologi penelitian tradisional.

6. Bagi Instansi dan Organisasi Non-Pemerintah

Organisasi yang berkecimpung dalam keberlanjutan dan mitigasi risiko bencana dapat memanfaatkan hasil penelitian ini untuk merancang program yang lebih efektif