

BAB I PENDAHULUAN

A. Latar Belakang

Iklim merupakan kondisi rata-rata cuaca yang terjadi dalam jangka waktu yang relatif panjang pada suatu wilayah yang luas, sehingga memerlukan proses pemantauan dan proyeksi secara berkelanjutan. Kondisi iklim memiliki pengaruh yang besar terhadap kelangsungan hidup manusia [1]. Indonesia terletak di wilayah yang dilalui garis katulistiwa serta beriklim tropis, sehingga memiliki dua musim utama, yaitu musim kemarau dan musim hujan. Seiring dengan perubahan waktu, cuaca dapat berubah. Curah hujan merupakan faktor penting dalam perubahan cuaca. Curah hujan didefinisikan sebagai jumlah air hujan yang mencapai permukaan bumi dalam suatu periode waktu tertentu selama selang waktu pengukuran [2]. Tingkat curah hujan yang terlalu tinggi maupun terlalu rendah dapat menimbulkan berbagai dampak, baik yang bersifat positif dan negatif bagi kehidupan manusia. Curah hujan sendiri dipengaruhi oleh beberapa faktor, antara lain suhu udara, kelembapan, tekanan udara, kecepatan angin, dan lama penyinaran matahari.

Curah hujan merupakan salah satu informasi yang sangat penting, khususnya pada sektor pertanian. Informasi curah hujan salah satunya berguna dalam mengantisipasi kemungkinan terjadinya kejadian ekstrem, yang dapat mengakibatkan produksi pertanian mengalami kegagalan. Data harian memungkinkan identifikasi awal musim hujan secara presisi. Bagi petani lahan tadah hujan, luaran ini penting untuk menentukan waktu yang optimal guna menghindari resiko kekeringan karena lama curah hujan rendah pada awal pertumbuhan atau banjir karena lama curah hujan tinggi pada saat masa panen.

Berdasarkan data dari Badan Pusat Statistika (BPS) provinsi Jawa Tengah, pada tahun 2021 Kabupaten Cilacap tercatat sebagai kabupaten dengan wilayah terluas di Provinsi Jawa Tengah, dengan luas sekitar 225.361 Hektare (Ha) atau 2.124,50 km², yang setara dengan 6,94% dari total wilayah Jawa Tengah [3]. Secara geografis, wilayah Kabupaten Cilacap terletak pada koordinat 108°4'30" - 109°30'30" Bujur Timur (BT) dan 7°30'00" - 7°45'20" Lintang

Selatan (LS) [3]. Kabupaten Cilacap juga dikenal sebagai daerah dengan potensi lahan pangan yang cukup luas. Data dari Badan Pusat Statistika (BPS) Provinsi Jawa Tengah, menunjukkan bahwa pada tahun 2021 luas lahan sawah di Kabupaten Cilacap mencapai 64,72 Hektare (Ha) atau sekitar 30,27% dari total luas wilayahnya [3]. Pada tahun yang sama, produksi padi sawah di Kabupaten Cilacap mencapai 838.758 Ton, meningkat 15,96% dibandingkan dengan produksi padi pada tahun 2019 [3]. Tingginya produksi padi tersebut menjadikan Kabupaten Cilacap sebagai salah satu daerah penopang utama ketahanan pangan di Provinsi Jawa Tengah [3].

Kabupaten Cilacap merupakan salah satu wilayah di Provinsi Jawa Tengah yang terdampak oleh fluktuasi tingkat curah hujan. Curah hujan yang terlalu tinggi akan berpotensi menyebabkan terjadinya banjir di beberapa wilayah. Berdasarkan informasi dari DISKOMINFO Kabupaten Cilacap salah satu kasus banjir akibat hujan turun dengan lama tinggi serta kondisi drainase yang tidak dapat menampung air pada tanggal 27 April 2023 yang mengakibatkan 8 kelurahan di Kecamatan Cilacap Selatan, Cilacap Tengah dan Cilacap Utara terdampak banjir dengan genangan antara 60 cm hingga 1,5 meter [4]. Begitu pula pada saat curah hujan terlalu rendah yang akan berakibat kekeringan di beberapa daerah juga. Berdasarkan informasi dari website pemerintah Kabupaten Cilacap salah satu kasus kekeringan akibat lama curah hujan yang rendah pada bulan September 2023 yang mengakibatkan kurang lebih 16 wilayah di Kabupaten Cilacap mengalami kekeringan, beberapa diantaranya meliputi wilayah Kawunganten, Patimuhan dan Bantarsari [5]. Oleh karena itu, diperlukan upaya prediksi curah hujan guna meminimalkan potensi dampak negatif yang dapat terjadi. Pola data curah hujan yang tidak menentu, menyebabkan prediksi manual menjadi memiliki keterbatasan. Pada umumnya, prediksi curah hujan dilakukan pada data histori yang kemudian dianalisis dengan menggunakan metode atau teknik tertentu.

Berbagai metode prediksi dapat digunakan untuk memperkirakan nilai pada data runtun waktu (*time series*). Pola yang tidak menentu yang dimiliki curah hujan juga akan menyebabkan kesulitan dalam memprediksi secara manual [6]. Oleh karena itu, pendekatan prediksi yang hanya didasarkan pada

metode statistika/konsep-konsep matematika memiliki keterbatasan. Akan tetapi, perkembangan *data mining* memungkinkan mesin untuk mempelajari pola data yang kompleks [7]. Selain itu, pendekatan lain yang dapat digunakan adalah metode *heuristik* modern yang didasarkan pada algoritma dalam bidang *soft computasi* (SC). Salah satu metode yang termasuk dalam pendekatan metode *heuristik* modern tersebut adalah metode *K-Nearest Neighbor* (KNN). Metode *K-Nearest Neighbor* (KNN) merupakan bagian dari teknik *data mining* yang mampu mengidentifikasi suatu pola data *sequensial*, dimana mempunyai karakteristik yaitu data diproses berdasarkan urutan waktu [6]. Dengan kemampuan tersebut, metode *K-Nearest Neighbor* (KNN) dapat dimanfaatkan sebagai salah satu teknik dalam *data mining* untuk melakukan proses prediksi [8].

Metode *K-Nearest Neighbor* (KNN) merupakan metode klasifikasi yang dapat menghasilkan sebuah prediksi, dimana didasarkan pada asumsi bahwa suatu objek dengan jarak yang berdekatan satu dengan yang lainnya juga akan cenderung mempunyai karakteristik yang serupa. Dengan demikian, apabila karakteristik suatu objek diketahui, maka objek lainnya juga akan dapat diprediksi berdasarkan tetangga terdekatnya [9]. Metode *K-Nearest Neighbor* (KNN) mengklasifikasikan suatu pengamatan yang belum diketahui ke dalam kelas mayoritas diantara pengamatan tetangga terdekatnya, yang ditentukan melalui perhitungan matrik jarak pada data latih. *K-Nearest Neighbor* (KNN) termasuk salah satu algoritma *Machine Learning* yang relatif sederhana dan dapat diterapkan dalam analisis data dengan dimensi variabel yang besar [10]. *Machine Learning* merupakan pendekatan komputasi cerdas yang memiliki potensi besar dalam bidang klasifikasi dan prediksi [11]. Prinsip kerja algoritma KNN adalah dengan mencari jarak terdekat antara data yang akan dievaluasi dengan sejumlah k tetangga terdekatnya dalam data latih [9]. Nilai k tersebut menunjukkan jumlah tetangga terdekat yang digunakan sebagai dasar dalam menentukan kategori prediksi kategori pada data *testing* [12]. Selain itu, penentuan variabel yang tepat sangat berpengaruh terhadap hasil dan tingkat akurasi yang dihasilkan [9]. Dalam proses prediksi curah hujan, beberapa variabel yang umum digunakan meliputi suhu udara, kelembapan,

takanan udara, kecepatan angin, dan lama penyinaran matahari. Variabel tersebut dapat dimanfaatkan untuk mengidentifikasi kecenderungan terjadinya hujan pada periode yang akan datang [6].

Pada penelitian sebelumnya yang dilakukan oleh Vita Fatimah, Ahmad Zaenal Arifin, dan Suzatmo Putro (2024), hasil prediksi menunjukkan nilai RMSE sebesar 36,3942. Nilai RMSE tersebut dipengaruhi oleh keterbatasan jumlah dataset yang digunakan dalam meramalkan curah hujan [13]. Penelitian lain dilakukan oleh Muhammad Yusuf R. R., Muhammad Valensyah A., dan Wawan Gunawan (2021) menggunakan sebanyak 3623 data dengan perbandingan data *training* dan data *testing* adalah 80 : 20, menghasilkan suatu prediksi dengan tingkat akurasi data sebesar 89%, dengan tingkat akurasi tersebut metode yang digunakan diharapkan dapat membantu dalam memprediksi suhu dan cuaca di Indonesia [14]. Selain itu, penelitian yang dilakukan oleh Deden Martia Nanda, Tacbir Hendro Pudjiantoro dan Puspita Nurul Sabrina (2022) menggunakan variabel temperatur rata-rata, kelembapan rata-rata, kecepatan angin rata-rata, dan curah hujan menghasilkan akurasi berdasarkan nilai $k = 5$ didapatkan hasil pengujian dengan *K-Cross Validation* maka nilai akurasi sebesar 86,199%. Sedangkan hasil pengujian dengan *Confusion Matrix* menghasilkan akurasi sebesar 85,78% [6].

Berdasarkan beberapa penelitian diatas, pembeda yang ada pada penelitian yang akan dilakukan adalah terletak pada lokasi penelitian yaitu penelitian dilakukan di Kabupaten Cilacap. Kemudian pada variabel yang digunakan yaitu adanya penambahan variabel, sehingga variabel yang digunakan pada penelitian ini adalah curah hujan harian, suhu udara rata-rata, kelembapan rata-rata, tekanan udara rata-rata, kecepatan angin rata-rata dan lama penyinaran matahari. Selanjutnya adalah alat analisis data yang digunakan, dimana penelitian ini menggunakan Python.

Berdasarkan latar belakang yang telah diuraikan diatas, maka diusulkan sebuah penelitian dengan judul **"Prediksi Curah Hujan Harian Dengan Metode *K-Nearest Neighbor* (KNN) Menggunakan Python (Studi Kasus di Stasiun Meteorologi Tunggul Wulung Cilacap)"**.

B. Rumusan Masalah

Berdasarkan uraian latar belakang diatas, permasalahan yang menjadi fokus penelitian ini dirumuskan sebagai berikut:

1. Bagaimana prediksi curah hujan harian dengan metode *K-Nearest Neighbor* (KNN)?
2. Bagaimana akurasi prediksi curah hujan harian dengan metode *K-Nearest Neighbor* (KNN)?

C. Batasan Masalah

Untuk memperoleh perumusan masalah yang lebih terarah, penelitian ini menetapkan batasan-batasan, antara lain sebagai berikut:

1. Data deret waktu (*time series*) yang digunakan dalam penelitian ini merupakan data harian yang mencakup curah hujan, suhu udara rata-rata, kelembapan rata-rata, tekanan udara rata-rata, kecepatan angin rata-rata, dan lama penyinaran matahari untuk periode Januari 2021 sampai dengan Desember 2025.
2. Menggunakan acuan 80:20 dari enam variabel untuk perbandingan data *training* dan data *testing*.

D. Tujuan Penelitian

Berdasarkan permasalahan yang telah diidentifikasi, tujuan dari penelitian ini dapat dirumuskan sebagai berikut:

1. Untuk mengetahui prediksi curah hujan harian dengan metode *K-Nearest Neighbor* (KNN).
2. Untuk mengetahui akurasi prediksi curah hujan harian dengan metode *K-Nearest Neighbor* (KNN).

E. Manfaat Penelitian

Adapun manfaat yang dihasilkan dari penelitian ini dapat dirinci sebagai berikut:

1. Manfaat Teoritis
 - a. Penelitian ini dapat digunakan sebagai wadah untuk menambah informasi serta wawasan pengetahuan yang berkaitan dengan prediksi curah hujan harian metode *K-Nearest Neighbor* (KNN).

- b. Penelitian ini dapat dijadikan sebagai acuan atau referensi bagi penelitian-penelitian selanjutnya.

2. Manfaat Praktis

a. Untuk Peneliti

Penelitian ini sebagai bentuk partisipasi aktif untuk memberikan kontribusi pemikiran dalam bidang keilmuan matematika.

b. Untuk Instansi

Penelitian ini dapat berguna bagi instansi, yang dalam hal ini adalah Stasiun Meteorologi Tunggul Wulung Cilacap sebagai informasi terkait penelitian curah hujan harian.

c. Untuk Masyarakat Umum

Penelitian ini dapat berguna untuk memberikan peringatan dini terhadap bencana hidrometeorologi, sehingga masyarakat dapat melakukan langkah antisipasi lebih awal. Pada sektor pertanian, informasi terkait prediksi curah hujan dapat dimanfaatkan untuk menentukan waktu yang tepat dalam penanaman, pemupukan, dan panen sehingga hasil produksi dapat dimaksimalkan.