

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Keadaan atmosfer dalam jangka pendek, atau cuaca, ditentukan oleh berbagai faktor seperti kecepatan angin, suhu, tekanan udara, hujan, dan fenomena atmosfer lainnya. Meskipun sulit memprediksi dengan pasti, curah hujan bisa diperkirakan menggunakan data historis. Contohnya, informasi mengenai curah hujan masa lalu dapat digunakan untuk meramalkan curah hujan di masa mendatang. Selain itu, parameter laut juga sangat berpengaruh dalam memprediksi curah hujan. (Subhan, 2017)

Studi yang dilakukan oleh Pai, Maya L., dan rekan-rekannya menekankan bagaimana parameter lautan dapat mempengaruhi prediksi curah hujan monsun barat daya dengan menggunakan jaringan syaraf tiruan (JST). Temuan dari penelitian tersebut menunjukkan bahwa pendekatan JST terbukti efektif dalam meramalkan curah hujan. Selain Jaringan Saraf Tiruan (JST), K-Nearest Neighbor (K-NN) juga berperan penting dalam menganalisis data dalam penelitian. (Sidik & Sen, 2019).

Saat merancang sistem untuk mengklasifikasikan kondisi cuaca seperti hujan ringan, hujan deras, atau hujan sangat deras, Anda dapat menerapkan metode K-Nearest Neighbor (K-NN) dengan menganalisis data cuaca yang dijadikan acuan dalam penelitian. Data mining juga berperan penting dalam memprediksi variabel dari data warehouse dan merupakan bagian dari proses penemuan *Database Knowledge Discovery in Databases* (KDD). Ini dapat digunakan di berbagai bidang seperti kecerdasan buatan dan *database*. Di bidang kesehatan, algoritma data mining digunakan untuk mengekstrak informasi dari *database* medis untuk membantu diagnosis penyakit. (Saputri, 2021).

Data mining melibatkan proses komputasi, teknik statistik, serta penggunaan clustering dan klasifikasi untuk mengidentifikasi pola dalam dataset. Tujuannya adalah mengekstrak informasi yang berguna dari dataset besar agar dapat dimengerti dan digunakan di masa depan (Hasanah et al., 2021).

Konsep *K-nearest neighbour* digunakan untuk mengelompokkan objek berdasarkan nilai K dalam suatu ruang fitur. Prosesnya memerlukan pengukuran jarak untuk menentukan kedekatan objek. Saat memproses data uji, objeknya dikelompokkan berdasarkan objek terdekatnya. Metode ini mengidentifikasi kelas dari tetangga terdekat dan memilih kelas yang paling umum di dalam kelompok tersebut. Pendekatan *K-nearest neighbour* mengambil keputusan klasifikasi berdasarkan pola yang ada dalam data latihan, dan kinerjanya bisa optimal ketika digunakan dengan kumpulan data yang besar.(Dinata & Hasdyna, 2020).

Penerapan data mining untuk mengetahui kemungkinan terjadinya hujan menggunakan algoritma *K-Nearest Neighbor* (K-NN) memungkinkan dilakukannya klasifikasi sehingga memudahkan masyarakat dalam menerima informasi yang simpang siur mengenai kemungkinan terjadinya hujan.

## 1.2 Rumusan Masalah

Masalah yang dirumuskan penulis berdasarkan latar belakang diatas adalah sebagai berikut :

1. Bagaimana merancang sebuah aplikasi yang dapat menentukan potensi Curah hujan?
2. Bagaimana cara menerapkan Metode algoritma *K-Nearest Neighbor* (KNN)?
3. Bagaimana Hasil Yang didapat dari pengujian menggunakan Data Mining dalam penelitian ini?

## 1.3 Tujuan Penelitian

Tujuan penelitian dalam tugas akhir ini adalah :

1. Mengetahui cara merancang sebuah aplikasi yang dapat menentukan potensi Curah hujan.
2. Memahami cara menerapkan Metode algoritma *K-Nearest Neighbor* (KNN).
3. Mengetahui hasil Pengujian Menggunakan Data Mining.

#### 1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah :

1. Membantu dan memudahkan masyarakat dalam mengetahui informasi potensi Curah hujan.
2. Sebagai wahana dalam mengembangkan kemampuan dan sebagai pengetahuan dalam mengembangkan ilmu Komputer, algoritma dan pemograman berbasis web.
3. Menambah ilmu pengetahuan bagi sang penulis dalam pemograman web, khususnya dalam penerapan Metode Algoritma *K Nearest Neighbor* (KNN).

#### 1.5 Ruang Lingkup dan Batasan Masalah

Batasan masalah pada sistem yang akan kita buat adalah :

1. Sistem dirancang untuk mengetahui potensi curah hujan dengan menggunakan algoritma K-Nearest Neighbors (KNN).
2. Kategori yang dapat dijadikan variabel dalam suatu penelitian adalah hujan berawan, hujan ringan, hujan sedang, hujan lebat, hujan sangat lebat, dan hujan ekstrim.
3. Jumlah kriteria yang digunakan dalam penelitian ini adalah 21 kriteria yaitu, tekanan udara (QFE), suhu udara bola kering T(\$BK), suhu udara bola basah T(\$BK), titik didih (Td), kelembapan (RH), penguapan udara (Ew), jumlah awan yang menutupi horitok langit (N), awan cumulonimbus (Cb), awan comulis (Cu), awan strato cumulus (Sc), temperature minimum (Tn), temperature maxximum (Tx), temperature rata-rata (Tavg), kelembapan rata-rata (RH\_avg), lamanya penyinaran matahari(Ss), kecepatan angin maksimum (ff\_x), arah angin saat kecepatan maksimum(ddd\_x), kecepatan angin rata-rata (ff\_avg), arah angin terbanyak (ddd\_car), curah hujan (RR), hujan (RA).
4. Data yang dikumpulkan adalah data tahun lalu di BMKG (Badan Meteorologi Klimatologi dan Geofisika) wilayah Aceh Utara. Data yang digunakan sebanyak 365 data, 265 data latih dan 100 data uji.

