

BAB I

PENDAHULUAN

1.1 Latar Belakang

Setiap hari kondisi iklim dan cuaca bisa berubah-ubah karena dipengaruhi oleh banyak faktor. Faktor penyebab perubahan cuaca antara lain suhu, kelembaban, angin, lama penyinaran matahari dan sebagainya. Banyak pekerjaan yang erat kaitannya dengan cuaca, sehingga cuaca sangat penting untuk kehidupan manusia (Rofiq et al., 2020). Kegiatan yang berkaitan erat dengan cuaca seperti pada sektor pertanian, perkebunan dan pada penerbangan (Hamami & Dahlan, 2022). Pada penelitian (Rofiq et al., 2020) disebutkan bahwa sektor pariwisata, kehutanan, konstruksi, pelayaran, kesehatan merupakan bidang pekerjaan yang juga membutuhkan informasi cuaca dalam proses kerjanya. Dengan adanya kondisi tersebut, perlu adanya suatu sistem yang dapat menunjukkan cuaca di hari yang akan datang untuk meminimalkan tingkat risiko dan dampak yang akan terjadi (Sasake et al., 2021).

Cuaca merupakan suatu kondisi udara yang bisa diamati dan terjadi dalam waktu yang singkat, cuaca terjadi pada waktu dan di wilayah tertentu (Luthfiarta et al., 2020). Cuaca terjadi pada waktu yang singkat normalnya hanya beberapa waktu saja. Misalkan kondisi udara pada waktu pagi bisa berubah pada waktu siang, kondisi udara siang bisa berubah pada waktu sore (Miftahuddin, 2016). Di Indonesia, Badan Meteorologi Klimatologi dan Geofisika (BMKG) menerbitkan prakiraan cuaca pada setiap harinya (Luthfiarta et al., 2020). Selain itu, mengeluarkan peringatan dini terhadap adanya bencana dan berhubungan dengan cuaca (Fatihin et al., 2020).

Indonesia memiliki dua musim yaitu musim kemarau dan musim penghujan. Musim kemarau berlangsung pada antara bulan Mei-Juni dan puncaknya pada bulan Juli-Agustus (Rangkuti et al., 2021). Sedangkan musim penghujan berlangsung pada bulan September hingga Maret (Badan Meteorologi Klimatologi dan Geofisika, 2022). Hujan ditentukan oleh besaran intensitas hujan pada suatu tempat yang mana terdapat siklus hidrologi dan unsur iklim (Fransiska et al., 2019). Parameter hujan yang dapat diukur disebut dengan curah hujan yang mana curah

hujan menunjukkan berapa banyak air yang jatuh pada suatu daerah (Fransiska et al., 2019). Banyaknya curah hujan tidak bisa ditentukan secara aktual, akan tetapi banyaknya curah hujan bisa diprediksi berdasarkan riwayat data yang telah ada (Oktaviani & Afdal, 2013). Jika tidak adanya suatu sistem yang dapat memperkirakan hujan di masa yang akan mendatang akan berdampak buruk bagi sejumlah sektor yang berkaitan langsung dengan cuaca.

Pada sektor pariwisata, berkaitan langsung dengan cuaca sehingga jika cuaca buruk maka akan mengubah preferensi wisatawan untuk mengunjungi tempat wisata. Hubungan antara cuaca dengan pariwisata yaitu melihat dari segi kesiapan wisatawan untuk mempersiapkan apa saja ketika cuaca sedang hujan atau ketika sedang terik. Dan melihat dari segi tempat kunjungan wisata, terdapat perubahan preferensi atau tetap berkunjung di tempat wisata (Lusiani & Wally, 2018). Sehingga penting adanya klasifikasi cuaca yang dapat memberikan informasi mengenai cuaca akan panas ataukah akan hujan sehingga dalam melakukan aktivitas bisa mempersiapkan diri dan dapat meminimalkan risiko yang terjadi.

Kabupaten Tuban merupakan salah satu kabupaten yang ada di Jawa Timur yang berbatasan langsung dengan laut jawa pada sebelah utara, sebelah selatan berbatasan langsung dengan Kabupaten Bojonegoro, sebelah barat berbatasan dengan Kabupaten Rembang dan sebelah timur berbatasan langsung dengan Kabupaten Lamongan. Dalam topografi wilayah, Kabupaten Tuban memiliki luas wilayah daratan, lautan dan pantai. Selain itu, wilayah tuban dibagi menjadi empat kawasan yaitu wilayah pantai, wilayah pertanian, wilayah aliran sungai dan wilayah pegunungan kapur (Ariana, 2016). Kabupaten Tuban memiliki banyak tempat wisata yang indah dan menarik. Seperti wisata air terjun, wisata religi, wisata sejarah dan lain-lain (Sovianingrum, 2019). Pada situs resmi Kabupaten Tuban (tubankab.go.id) disebutkan bahwa Tuban memiliki banyak tempat wisata, diantaranya Pantai Kelapa, Museum Kambang Putih, Pantai Boom. Selain itu juga terdapat Gua Akbar, Makam Sunan Bonang, Pemandian Bektiharjo dan sebagainya. Dapat dilihat bahwasanya terdapat banyak tempat wisata baik di dalam ruangan maupun di luar ruangan terutama yang berhubungan langsung dengan alam. Dengan adanya kondisi tersebut, penentuan cuaca yang tepat merupakan hal yang harus diperhatikan mengingat persiapan-persiapan yang dibutuhkan ketika hujan

atau ketika cuaca sedang terik. Sehingga dapat meminimalkan risiko yang akan terjadi. Dengan demikian, penulis memiliki gagasan untuk mengembangkan sistem yang di dalamnya berisikan klasifikasi cuaca berbasis fitur penentu hujan pada Kabupaten Tuban.

Metode klasifikasi seperti *Decision Tree*, *Naïve Bayes*, *K-Nearest Neighbor* digunakan klasifikasi cuaca. *Decision Tree* merupakan salah satu metode klasifikasi yang mudah dipahami (Lakshmi et al., 2016). Bentuk representasi dari metode ini dengan menggunakan hierarki pohon (*tree*) yang di dalamnya terdapat sebuah *node* dan *leaf* (Permana et al., 2021). Kelebihan dari metode ini bersifat fleksibel yang dapat meningkatkan akurasi yang dihasilkan. Selain itu, mempersingkat atribut pada setiap node tanpa banyak mengurangi akurasi yang dihasilkan. Disamping itu, metode ini terjadi *overlap* jika penggunaan kelas dan kriterianya banyak (Dolok et al., 2021).

Naïve bayes merupakan salah satu metode klasifikasi yang menganut teorema bayes yang mana pada penerapannya setiap atribut bersifat bebas tidak memiliki korelasi atau ketergantungan (Lukito & Chrismanto, 2015). Atribut yang independent ini dapat memengaruhi tingkat akurasi yang dihasilkan, namun proses perhitungan dari metode cepat dan efisien karena memiliki perhitungan yang sederhana (Permana et al., 2021)

Menurut (Khairi, Ghozali, & Nur, 2021) *K-Nearest Neighbor* merupakan algoritma yang dalam proses klasifikasinya berdasarkan nilai tetangga terdekatnya atau biasa disebut *k*, selain digunakan untuk klasifikasi *K-Nearest Neighbor* dapat juga digunakan untuk prediksi. Untuk memecahkan masalah klasifikasi algoritma *K-Nearest Neighbor* menghasilkan hasil yang kompetitif dan signifikan dan paling sederhana (Nikmatun et al., 2019). Selain itu, kelebihan dari *K-Nearest Neighbor* yaitu performa yang tetap bagus meskipun ada data yang *noise* (Devita et al., 2018). Sedangkan kekurangan dari algoritma *K-Nearest Neighbor* yaitu penurunan akurasi yang disebabkan pengaruh skala input dan pengukuran jarak *Euclidean* yang memperlakukan atribut data secara sama yang tidak sesuai dengan data yang sebenarnya (Hocke & Martinetz, 2013).

Beberapa penelitian terdahulu yang telah dilakukan oleh Nursobah, Siti Lailiyah, Bartolomius Harpad, Muhammad Fahmi pada tahun 2022 didapatkan

hasil penelitian bahwasanya algoritma *K-Nearest Neighbor* dapat menjadi model untuk proses penyelesaian pada permasalahan prediksi perkiraan hujan (Nursobah et al., 2022). Penelitian lainnya dilakukan oleh Arwansyah dan John S. Arie pada tahun 2019 mengenai prediksi curah hujan dan temperatur untuk tanaman padi menggunakan algoritma *K-Nearest Neighbor* menghasilkan informasi yang tepat sehingga dapat dikatakan penggunaan algoritma *K-Nearest Neighbor* telah berhasil menyelesaikan permasalahan tersebut (Arwansyah & Arie, 2019).

Penelitian terdahulu juga dilakukan oleh Muhammad Naja Maskuri, Harliana, Kadek Sukerti dan R.M. Herdian Bhakti pada tahun 2022 yang meneliti tentang penerapan algoritma *K-Nearest Neighbor* untuk memprediksi penyakit stroke dengan hasil akhir penelitian mendapatkan nilai akurasi sebesar 95% dengan nilai $k=9$ (Maskuri et al., 2022). Pada tahun yang sama juga dilakukan penelitian oleh Arvin Christopher dan Teady Matius Surya Mulyana yang membahas mengenai klasifikasi tumbuhan angiospermae menggunakan algoritma *K-Nearest Neighbor* berdasarkan pada bentuk daun, hasil akhir penelitian ini mendapatkan nilai akurasi sebesar 81% (Christopher & Mulyana, 2022).

Berdasarkan latar belakang yang telah dipaparkan di atas, tujuan dilakukan penelitian ini untuk meminimalkan risiko akibat hujan dan dapat waspada akan cuaca yang akan mendatang, selain itu mengetahui bagaimana penerapan algoritma K-NN jika diterapkan pada data cuaca dan untuk mengetahui tingkat akurasi dari algoritma K-NN. Metode ini dipilih dibandingkan metode lainnya yang telah dipaparkan sebelumnya karena memiliki tingkat akurasi yang lebih tinggi dan kelemahannya bisa diatasi dengan memaksimalkan penghitungan *Euclidean Distance*.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, maka rumusan masalah pada penelitian ini yaitu sebagai berikut:

1. Bagaimana klasifikasi cuaca berbasis faktor penentu hujan menggunakan algoritma *K-Nearest Neighbor* (K-NN) pada wilayah Kabupaten Tuban?
2. Bagaimana evaluasi algoritma *K-Nearest Neighbor* (K-NN) algoritma pada klasifikasi cuaca berbasis faktor penentu hujan?

1.3 Batasan Masalah

Adapun Batasan masalah pada penelitian ini adalah:

1. Data yang digunakan bersumber dari dataonline.bmkg.go.id
2. Data yang digunakan hanya data mulai 01 Januari 2022 hingga 31 Januari 2023.
3. Lokasi yang digunakan hanya wilayah Kabupaten Tuban.
4. Pengembangan sistem berbasis *website*.

1.4 Tujuan Penelitian

Adapun tujuan yang dicapai pada penelitian ini yaitu sebagai berikut:

1. Untuk mengetahui klasifikasi cuaca berbasis faktor penentu hujan menggunakan algoritma *K-Nearest Neighbor* (K-NN) pada wilayah Kabupaten Tuban.
2. Untuk mengukur evaluasi algoritma *K-Nearest Neighbor* (K-NN) pada klasifikasi cuaca berbasis faktor penentu hujan.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini yaitu untuk:

1. Membantu memberikan kemudahan akses informasi untuk mengetahui cuaca yang akan datang.
2. Memberikan wawasan baru di bidang ilmu pengetahuan khususnya di bidang data mining klasifikasi.
3. Memberikan wawasan baru tentang penerapan algoritma *K-Nearest Neighbor* pada studi kasus klasifikasi cuaca.