

TUGAS AKHIR

**KLASIFIKASI CUACA DENGAN MENGGUNAKAN
ALGORITMA *K-NEAREST NEIGHBOR***



Oleh:

Dandy 1822250084

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN REKAYASA
UNIVERSITAS MULTI DATA PALEMBANG
PALEMBANG
2023**

**Fakultas Ilmu Komputer dan Rekayasa
Universitas Multi Data Palembang**

Program Studi Informatika
Tugas Akhir Sarjana Komputer
Semester Genap Tahun 2022/2023

**Klasifikasi Cuaca dengan Menggunakan Algoritma
*K-Nearest Neighbor***

Dandy 1822250084

Abstrak

Cuaca merupakan kejadian alam yang menjadi faktor penentu dalam aktivitas sehari-hari manusia. Cuaca ditentukan oleh berbagai macam faktor, beberapa diantaranya adalah tekanan, kecepatan angin, curah hujan, dan suhu udara. Faktor-faktor cuaca sangat berdampak bagi aktivitas manusia sehingga perlu dilakukan penelitian. Penelitian ini melakukan klasifikasi terhadap 3 kelas cuaca yaitu cerah (*sunny*), berawan (*cloudy*), dan hujan (*rainy*) menggunakan algoritma *K-Nearest Neighbor* sebagai algoritma klasifikasi cuaca dengan parameter nilai K sebesar 3, 5, 7, dan 9. Dataset cuaca yang hendak diteliti berupa 96453 baris data yang diambil dari website *Kaggle*. Dataset dilakukan pembagian atas data latih dan data uji dengan rasio 80:20. Implementasi algoritma *K-Nearest Neighbor* menghasilkan *confusion matrix* dan *classification report*. Hasil pengujian pada *confusion matrix* mengindikasikan jumlah data terprediksi benar terbanyak pada nilai K = 9 yaitu sebanyak 13132 data terprediksi benar. Adapun hasil pengujian pada *classification report* mengindikasikan tingkat akurasi tertinggi pada nilai K = 9 yakni sebesar 68,073%. Selain itu, pada *classification report*, semakin besar parameter nilai uji K, semakin besar nilai *accuracy*, *precision*, *recall*, dan *f1-score* baik pada kelas cuaca *cloudy*, *rainy*, maupun *sunny*.

Kata kunci: Cuaca, *K-Nearest Neighbor*, *Confusion Matrix*, *Classification Report*

BAB 1

PENDAHULUAN

Bab ini memaparkan garis besar fenomena penelitian yang hendak dikaji. Bab ini terdiri dari enam subbab, yakni latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian, dan sistematika penulisan.

1.1 Latar Belakang

Cuaca adalah kejadian pada alam berkenaan dengan kondisi atmosfer yang berlangsung di bumi dalam jangka waktu yang pendek dan menjadi faktor penentu aktivitas manusia dalam kehidupan sehari-hari. Cuaca terjadi karena reaksi suhu dan kelembaban yang berbeda antara satu tempat dengan tempat lainnya (Brawijaya, Felix, & Suryaningrum, 2020).

Cuaca ditentukan oleh berbagai macam faktor, beberapa diantaranya adalah keadaan atmosfer dalam kurun waktu yang pendek seperti tekanan, kecepatan angin, curah hujan, suhu, dan fenomena pada udara. Dengan berbagai macam faktor yang hendak diteliti, akan sangat sulit untuk membuat prakiraan cuaca yang akurat. Kesulitan tersebut disebabkan karena prakiraan cuaca pada kondisi dimana cuaca sukar untuk diprediksi menghasilkan hasil prakiraan cuaca yang tidak akurat. Hal ini mengakibatkan masyarakat kurang dapat memanfaatkan informasi prakiraan cuaca untuk mengantisipasi dampak yang ditimbulkan oleh cuaca sehingga dapat

mengganggu kegiatan sehari-hari masyarakat dalam berbagai macam sektor, beberapa diantaranya yakni sektor pertanian, perkebunan, dan penerbangan (Siregar, Tukino, Faisal, Fauzi, & Kadori, 2020).

Klasifikasi objek pada dataset tertentu dapat dilakukan dalam proses yang sangat sederhana dan mudah dengan menggunakan algoritma *K-Nearest Neighbor*. Alasan digunakannya algoritma *K-Nearest Neighbor* sebagai klasifikasi telah dikemukakan oleh banyak peneliti. Brawijaya dkk, (2020) menjelaskan bahwa algoritma *K-Nearest Neighbor* merupakan algoritma yang dapat diimplementasikan dengan baik karena terdapat sedikit variabel pada dataset yang digunakan dalam proses klasifikasi. Adapun Putra, Pardede, & Syahputra (2022) menjelaskan bahwa algoritma *K-Nearest Neighbor* merupakan metode klasifikasi objek pemalas yang memiliki performa yang sangat baik. Hal ini disebabkan oleh hasil klasifikasi dengan tingkat akurasi tinggi pada algoritma *K-Nearest Neighbor* karena mampu mengelompokkan objek baru berdasarkan data latih sebagai pembelajar dan data uji yang akan digunakan untuk proses klasifikasi.

Terdapat beberapa penelitian yang menjadi acuan pada penelitian berikut. Penelitian yang dilakukan oleh Siregar dkk, (2020) membahas mengenai Klasifikasi untuk Prediksi Cuaca Menggunakan *Ensemble Learning*. Klasifikasi prediksi menggunakan *ensemble learning* dengan hasil pengujian klasifikasi prediksi cuaca yang diperoleh berupa akurasi sebesar 81,21% dan MSE (*Means Square Error*) 18,79%.

Terdapat pula penelitian yang dilakukan oleh Lattifia, Buana, & Rusjyanthi (2022) yang membahas mengenai Model Prediksi Cuaca Menggunakan Metode LSTM. Prediksi cuaca dilakukan dengan menggunakan metode LSTM (*Long Short Term Memory*). Hasil pengujian pada model prediksi cuaca yang diperoleh berupa nilai RMSE dan MAPE terbaik sebesar 1,7444% dan 1,9499%.

Dan yang terakhir terdapat penelitian yang dilakukan oleh Luthfiarta, Febriyanto, Lestiawan, & Wicaksono (2020) yang membahas mengenai Analisa Prakiraan Cuaca dengan Parameter Suhu, Kelembaban, Tekanan Udara, dan Kecepatan Angin Menggunakan Regresi Linier Berganda. Prakiraan cuaca dilakukan dengan menggunakan pendekatan pemodelan *supervised learning* dengan regresi linier berganda sebagai algoritma estimasi. Hasil pengujian prakiraan cuaca yang diperoleh berupa nilai koefisien determinasi R² sebesar 25,5 persen.

Uraian di atas mengindikasikan bahwa penelitian terdahulu mengenai prediksi cuaca telah dilakukan dengan berbagai macam metode yang berbeda-beda. Maka dari itu, penelitian ini menggunakan algoritma *K-Nearest Neighbor* sebagai suatu pendekatan baru dalam melakukan klasifikasi cuaca dengan hasil klasifikasi cuaca berupa tiga kategori yakni cerah, berawan, dan hujan yang didasarkan pada faktor temperatur, kelembaban udara, tekanan udara, dan kecepatan angin. Melalui pembagian data latih dan data uji menggunakan algoritma *K-Nearest Neighbor*, penelitian ini hendak memperoleh hasil klasifikasi cuaca berdasarkan keempat faktor tersebut.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang telah dipaparkan, terciptalah rumusan masalah dalam penelitian ini yakni bagaimana cara melakukan klasifikasi cuaca dengan menggunakan metode *K-Nearest Neighbor*.

1.3 Ruang Lingkup

Dalam penelitian ini ditentukan ruang lingkup atau batasan masalah sebagai berikut:

1. Hasil klasifikasi cuaca yang akan diklasifikasi terbagi ke dalam 3 kelas cuaca yakni berawan (*cloudy*), hujan (*rainy*), dan cerah (*sunny*) pada variabel *summary* (rangkuman cuaca) sebagai data uji serta variabel lain berupa suhu, kelembaban, kecepatan angin, dan tekanan udara sebagai data latih dimana dataset tersebut berjumlah sebanyak 96453 data untuk masing-masing variabel.
2. Sampel dataset berkenaan dengan penelitian mengenai klasifikasi cuaca merupakan sampel dataset cuaca yang diambil dari dataset *Weather Dataset* yang dapat diakses pada <https://www.kaggle.com/datasets/muthuj7/weather-dataset>.
3. Dataset dibagi menjadi 80% data latih dan 20% data uji.
4. Dataset yang akan digunakan dalam penelitian ini adalah berupa bentuk format csv (*comma separated value*).
5. Metode klasifikasi yang digunakan adalah algoritma *K-Nearest Neighbor* (KNN) sebagai algoritma pengujian klasifikasi cuaca dengan nilai k yang digunakan adalah 3, 5, 7, dan 9.

6. Implementasi algoritma *K-Nearest Neighbor* dilakukan dengan bantuan aplikasi pemrograman *Google Colaboratory* dalam proses pengolahan dataset serta klasifikasi cuaca.
7. Bahasa pemrograman yang digunakan pada implementasi algoritma *K-Nearest Neighbor* dalam pelatihan dan pengujian data serta klasifikasi cuaca adalah bahasa pemrograman *Python*.

1.4 Tujuan dan Manfaat

Penelitian ini bertujuan untuk mengimplementasikan dan mengetahui hasil klasifikasi dari penggunaan algoritma *K-Nearest Neighbor* untuk klasifikasi cuaca. Adapun manfaat atau kegunaan dari penelitian ini yakni memahami cara mengimplementasikan algoritma *K-Nearest Neighbor* dalam klasifikasi cuaca guna memperkaya pengetahuan manusia dalam klasifikasi cuaca serta menambah referensi yang bermanfaat terkait penelitian algoritma *K-Nearest Neighbor*.

1.5 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir terdiri dari lima bab, yakni pendahuluan, landasan teori, metode penelitian, hasil dan pembahasan, serta simpulan dan saran.

BAB 1 PENDAHULUAN

Bab ini menjelaskan mengenai garis besar laporan tugas akhir melalui pemaparan latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan teori-teori yang mendukung dalam penulisan laporan tugas akhir terkait dengan cuaca dan implementasi algoritma *K-Nearest Neighbor* sebagai algoritma klasifikasi cuaca.

BAB 3 METODE PENELITIAN

Bab ini menjelaskan mengenai kebutuhan perangkat serta cara-cara mengimplementasikan algoritma *K-Nearest Neighbor*.

BAB 4 HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil pengujian sistem beserta kajian analisis dari hasil pengujian pada penelitian.

BAB 5 PENUTUP

Bab ini memaparkan mengenai rangkuman hasil penelitian pada laporan tugas akhir beserta saran pengembangan penelitian.

DAFTAR PUSTAKA

- Amalia, A., Zaidiah, A., & Isnainiyah, I. N. (2022). Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika*, 496-507.
- Atmaja, D. M., Hakim, A. R., Haryadi, D., & Suwaryo, N. (2021). Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Pengelompokan Tingkat Risiko Penyebaran Covid-19 Jawa Barat. *Prosiding Seminar Nasional Teknologi Energi dan Mineral*, 1218-1226.
- Brawijaya, J., Felix, & Suryaningrum, K. M. (2020). Aplikasi Pendeteksi dan Analisa Cuaca Menggunakan Metode K-Nearest Neighbor Berbasis Android. *Jurnal Sarjana Teknik Informatika*, 1-10.
- Farokhah, L. (2020). Implementasi K-Nearest Neighbor Untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 1129-1136.
- Hardiyanto, B., & Rozi, F. (2020). Prediksi Penjualan Sepatu Menggunakan Metode K-Nearest Neighbor. *Journal of Education and Information Communication Technology*, 13-18.
- Hidayati, N., & Hermawan, A. (2021). K-Nearest Neighbor (K-NN) Algorithm with Euclidean and Manhattan in Classification of Student Graduation. *Journal of Engineering and Applied Technology*, 86-91.
- Id, I. D. (2021). *Machine Learning : Teori, Studi Kasus, dan Implementasi Menggunakan Python*. Pekanbaru: Badan Penerbit Universitas Riau.
- Isnain, A. R., Supriyanto, J., & Kharisma, M. P. (2021). Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning. *Indonesian Journal of Computing and Cybernetics Systems*, 121-130.
- Istighfarizky, F., ER, N. A., Widiartha, I. M., Astuti, L. G., Putra, I. G., & Suhartana, I. K. (2022). Klasifikasi Jurnal menggunakan Metode KNN dengan

Mengimplementasikan Perbandingan Seleksi Fitur. *Jurnal Elektronik Ilmu Komputer Udayana*, 167-176.

Lattifia, T., Buana, P. W., & Rusjyanthi, N. D. (2022). Model Prediksi Cuaca Menggunakan Metode LSTM. *Jurnal Ilmiah Teknologi dan Komputer*, 994-1000.

Luthfiarta, A., Febriyanto, A., Lestiawan, H., & Wicaksono, W. (2020). Analisa Prakiraan Cuaca dengan Parameter Suhu, Kelembaban, Tekanan Udara, dan Kecepatan Angin Menggunakan Regresi Linear Berganda. *Journal of Information System*, 10-17.

Muhammad, D. I., Ermatita, & Falih, N. (2021). Penggunaan K-Nearest Neighbor (KNN) untuk Mengklasifikasikan Citra Belimbing Berdasarkan Fitur Warna. *Jurnal Informatik*, 9-16.

Mustafa, M. S., & Simpen, I. W. (2019). Implementasi Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Pasien Terkena Penyakit Diabetes Pada Puskesmas Manyampa Kabupaten Bulukumba. *Prosiding Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 1-10.

Purwaningsih, E., & Nurelasari, E. (2021). Penerapan K-Nearest Neighbor Untuk Klasifikasi Tingkat Kelulusan Pada Siswa. *Syntax: Jurnal Informatika*, 46-56.

Putra, P., Pardede, A. M., & Syahputra, S. (2022). Analisis Metode K-Nearest Neighbor (KNN) dalam Klasifikasi Data Iris Bunga. *Jurnal Teknik Informatika Kaputama*, 297-305.

Setiawan, M. R., Sari, Y. A., & Adikara, P. P. (2019). Klasifikasi Citra Makanan Menggunakan K-Nearest Neighbor dengan Fitur Bentuk Simple Morphological Shape Descriptors dan Fitur Warna Grayscale Histogram. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2726-2731.

Siregar, A. M., Tukino, Faisal, S., Fauzi, A., & Kadori, I. (2020). Klasifikasi untuk Prediksi Cuaca Menggunakan Ensemble Learning. *Jurnal Pengkajian dan Penerapan Teknik Informatika*, 138-147.