

Implementasi Metode K–Nearest Neighbors Untuk Prediksi Cuaca Dengan Konsep Data Mining

Ismail^{*1}, Irma², Lira Sarwana³

^{1,2,3}Universitas Lamappapoleonro

^{1,2,3}Jl. Kesatria No.60 Watansoppeng, Soppeng Sulawesi Selatan

Email: ¹ismail@unipol.ac.id, ²irma@unipol.ac.id, ³lira@unipol.ac.id

(Naskah masuk: 14-09-2024, direvisi: 08-10-2024, diterbitkan: 29-11-2024)

Abstrak

Prediksi cuaca terus berkembang seiring dengan kemajuan dalam ilmu atmosfer, teknologi komputasi, dan analisis data. Namun, penting untuk diingat bahwa cuaca adalah sistem yang kompleks dan sulit untuk diprediksi dengan sempurna. Prediksi cuaca masih memiliki keterbatasan dan tingkat ketidakpastian tertentu, terutama dalam jangka waktu yang lebih lama dan dalam menghadapi peristiwa cuaca ekstrem yang kompleks. Data hasil pengamatan ini sangatlah penting untuk melihat karakteristik cuaca di tempat tersebut dan pembuatan informasi perkiraan cuaca beberapa hari ke depan. Setiap harinya Badan Meteorologi dan Geofisika (BMKG) melakukan pengukuran terhadap faktor-faktor cuaca. Ketersediaan data yang melimpah tentang faktor-faktor cuaca tersebut akan mendukung penggalian informasi untuk prakiraan cuaca, Sementara itu dalam proses pembuatan informasi perkiraan cuaca terdapat beberapa hambatan, Pertama susahnya membuat informasi perkiraan cuaca di karenakan melibatkan lebih dari satu sumber data seperti data hasil pengamatan, data hasil model aplikasi cuaca, data hasil gambar kondisi awan dari satelit, data hasil kondisi awan dari radar. Hambatan Kedua, perkiraan cuaca pada umumnya mengandalkan keahlian dari seorang prakirawan cuaca, sehingga interpretasi yang di hasilkan bisa berbeda antar prakirawan satu dengan yang lain karena bergantung dari pengalaman dan keahlian masing-masing. Dengan menggabungkan teknik data mining dengan pemahaman mendalam tentang ilmu atmosfer, para peneliti cuaca dapat menghasilkan prediksi cuaca yang lebih andal dan akurat. Ini memiliki implikasi penting dalam berbagai bidang, termasuk pertanian, transportasi, manajemen bencana, dan banyak lagi. Penerapan teknik data mining untuk sistem prediksi dapat menggunakan berbagai metode salah satunya adalah metode K-Nearest Neighbors. KNN bekerja dengan cara mencari k-tetangga terdekat dari data yang ingin diklasifikasikan atau diprediksi, lalu mengambil mayoritas kelas tetangga tersebut sebagai prediksi kelas atau nilai regresi untuk data tersebut. Hasil implementasi Sistem Analisis Pengujian Metode Knn (K–Nearest Neighbors) Pada Studi Kasus Prediksi Cuaca Menggunakan Model Data Mining berjalan dengan baik dan menghasilkan tiga bentuk prediksi berdasarkan prediksi yaitu hujan dan tidak hujan. Hasil implementasi mendapatkan 33% Prediksi hujan, dan 75% Prediksi tidak hujan dari total data cuaca sebanyak 8 Data.

Kata kunci: K–Nearest Neighbors, Prediksi Cuaca, Data Mining.

Abstract

Weather prediction continues to evolve along with advances in atmospheric science, computing technology, and data analysis. However, it is important to remember that weather is a complex system and difficult to predict perfectly. Weather prediction still has limitations and a certain degree of uncertainty, especially over longer time frames and in the face of complex extreme weather events. This observation data is very important to see the characteristics of the weather in that place and make weather forecast information for the next few days. Every day the Meteorological and Geophysical Agency (BMKG) takes measurements of weather factors. The availability of abundant data on these weather factors will support the extraction of information for weather forecasts, Meanwhile, in the process of making weather forecast information there are several obstacles, First, the difficulty of making weather forecast information is because it involves more than one data source such as observation data, data from weather application models, data from cloud condition images from satellites, data from cloud condition results from radar. The second obstacle is that weather forecasts generally rely on the expertise of a weather forecaster, so that the resulting interpretation can be different from one forecaster to another because it depends on the experience and expertise of each. By combining data mining techniques with a deep understanding of atmospheric science, weather researchers can produce more reliable and accurate weather predictions. This has important implications in a variety of fields, including agriculture, transport, disaster management, and more. The application of data mining techniques for prediction systems can use various methods, one of which is the K-Nearest Neighbors method. KNN works by finding the k-nearest neighbours of the data to be classified or predicted, then taking the majority of the neighbouring classes as the predicted class or regression value for the data. The results of the implementation of the Knn (K-Nearest Neighbors) Method Testing Analysis System on

Weather Prediction Case Studies Using Data Mining Models run well and produce three forms of predictions based on predictions, namely rain and no rain. The implementation results get 33% rain prediction, and 75% non-rain prediction from a total of 8 weather data.

Keywords: *K-Nearest Neighbours, Weather Prediction, Data Mining.*

1. PENDAHULUAN

Cuaca adalah keadaan udara yang terjadi di tempat sempit dan berlangsung dalam waktu singkat. Kondisi cuaca di suatu tempat dapat ditentukan oleh banyak faktor, seperti tekanan udara, kelembaban, angin, penyinaran matahari, dan sebagainya. Besarnya curah hujan yang terjadi tidak dapat ditentukan secara pasti, namun dapat diprediksi atau diperkirakan. Dengan menggunakan data historis besarnya curah hujan beberapa waktu yang lampau maka dapat diprediksi berapa besarnya curah hujan yang terjadi pada masa yang akan datang (Gelinis et al., 2020).

Secara keseluruhan prediksi cuaca mencakup tantangan ilmiah dan teknis yang signifikan, serta dampak sosial dan ekonomi yang luas. Prediksi cuaca yang lebih baik memiliki potensi untuk meningkatkan kesiapan dan pengambilan keputusan manusia dalam menghadapi perubahan cuaca yang dinamis. Sehingga dengan melihat faktor-faktor ini dapat di perkirakan cuaca yang akan terjadi pada esok harinya (Setiyani et al., 2020).

Prediksi cuaca terus berkembang seiring dengan kemajuan dalam ilmu atmosfer, teknologi komputasi, dan analisis data. Namun, penting untuk diingat bahwa cuaca adalah sistem yang kompleks dan sulit untuk diprediksi dengan sempurna. Prediksi cuaca masih memiliki keterbatasan dan tingkat ketidakpastian tertentu, terutama dalam jangka waktu yang lebih lama dan dalam menghadapi peristiwa cuaca ekstrem yang kompleks (Perdana, 2022).

Dinas Pertanian Kabupaten Soppeng adalah salah satu lembaga yang ada di Kabupaten Soppeng yang melaksanakan pengamatan sinoptik dan sebagian di antaranya memberi pelayanan analisis dan perkiraan cuaca. Data hasil pengamatan ini sangatlah penting untuk melihat karakteristik cuaca di tempat tersebut dan pembuatan informasi perkiraan cuaca

beberapa hari ke depan. Setiap harinya Badan Meteorologi dan Geofisika (BMKG) melakukan pengukuran terhadap faktor-faktor cuaca. Ketersediaan data yang melimpah tentang faktor-faktor cuaca tersebut akan mendukung penggalian informasi untuk prakiraan cuaca, Sementara itu dalam proses pembuatan informasi perkiraan cuaca terdapat beberapa hambatan, Pertama sulitnya membuat informasi perkiraan cuaca di karenakan melibatkan lebih dari satu sumber data seperti data hasil pengamatan, data hasil model aplikasi cuaca, data hasil gambar kondisi awan dari satelit, data hasil kondisi awan dari radar (Tita Lattifia et al., 2022). Hambatan Kedua, perkiraan cuaca pada umumnya mengandalkan keahlian dari seorang prakirawan cuaca, sehingga interpretasi yang di dihasilkan bisa berbeda antar prakirawan satu dengan yang lain karena bergantung dari pengalaman dan keahlian masing-masing. Perbedaan interpretasi dapat membingungkan penerima informasi yang pada akhirnya berpeluang untuk menurunkan kualitas informasi yang disampaikan (Ulinuha & Farida, 2018). Untuk itu pentingnya melakukan pengembangan teknologi dalam melakukan sistem prediksi cuaca untuk membantu memberikan informasi yang akurat. Salah satu teknologi yang dapat diterapkan dalam memprediksi cuaca adalah sistem data mining.

Prediksi cuaca dengan menggunakan teknik data mining adalah aplikasi penting dalam bidang meteorologi dan ilmu atmosfer (Dwiyanti & Prianto, 2023). Data mining merupakan proses ekstraksi informasi yang berharga dan tersembunyi dari kumpulan data besar dengan menggunakan berbagai teknik analisis statistik, matematika, dan komputasi (Ismail Ismail, Rezky Erwin Syah, 2024). Dalam konteks prediksi cuaca, data mining digunakan untuk mengolah data historis cuaca dan informasi

lainnya guna memprediksi kondisi cuaca di masa depan.

Dengan menggabungkan teknik data mining dengan pemahaman mendalam tentang ilmu atmosfer, para peneliti cuaca dapat menghasilkan prediksi cuaca yang lebih andal dan akurat. Ini memiliki implikasi penting dalam berbagai bidang, termasuk pertanian, transportasi, manajemen bencana, dan banyak lagi. Penerapan teknik data mining untuk sistem prediksi dapat menggunakan berbagai metode salah satunya adalah metode K-Nearest Neighbors (KNN) (Rofiq et al., 2020).

Metode KNN adalah sebuah metode dalam analisis data yang digunakan untuk melakukan klasifikasi atau regresi. Metode ini mengandalkan prinsip bahwa data yang memiliki karakteristik serupa cenderung berada dalam lingkungan yang sama. KNN bekerja dengan cara mencari k-tetangga terdekat dari data yang ingin diklasifikasikan atau diprediksi, lalu mengambil mayoritas kelas tetangga tersebut sebagai prediksi kelas atau nilai regresi untuk data tersebut (Agung et al., 2023).

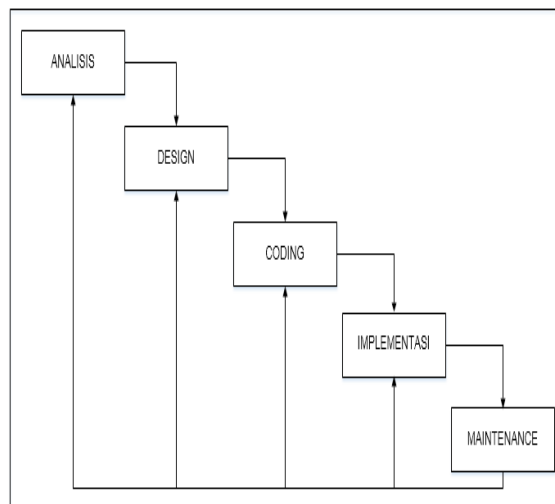
KNN memiliki beberapa kelebihan, seperti kemudahan implementasi dan kemampuan untuk menangani data yang kompleks atau non-linear. Namun, metode ini juga memiliki beberapa kelemahan, seperti sensitivitas terhadap data pencilan (outliers) dan kinerja yang buruk pada dataset dengan dimensi tinggi. Untuk itu penelitian yang akan dilakukan mencoba menerapkan metode KNN untuk menguji tingkat keakuratan data dalam sistem prediksi cuaca dengan model data mining. Gambaran dari sistem yang akan diterapkan yaitu melakukan sistem prediksi cuaca hujan atau tidak dengan parameter yang digunakan seperti temperatur, kecepatan angin, kelembaban udara (Purnamawati et al., 2022). Nilai dari setiap parameter akan dilakukan proses perhitungan algoritma dengan menggunakan KNN, hal ini sekaligus akan menampilkan tingkat akurasi kecocokan prediksi dari metode KNN (Udjulawa, 2023).

2. METODE PENELITIAN

Untuk mencapai hasil penelitian diperlukan metode penelitian yang jelas. Adapun metode tahapan penelitian sebagai berikut:

2.1 Tahapan Penelitian

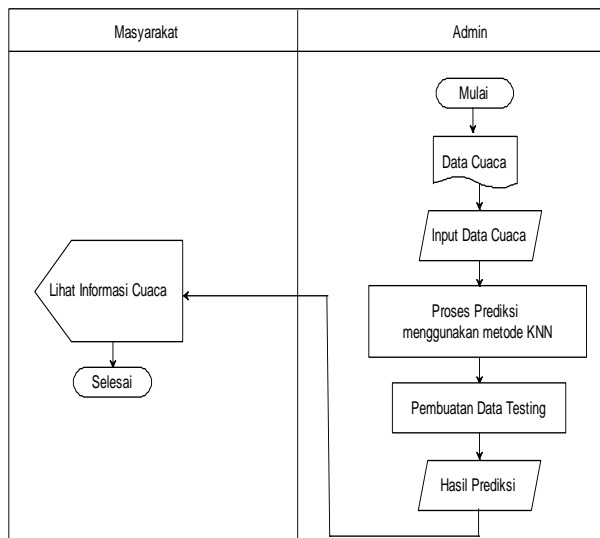
Penelitian ini menggunakan metodologi yang menerapkan metodologi model waterfall pada tahapan pengerjaan disertai dari awal hingga akhir. Model waterfall merupakan tahapan pekerjaan yang dilakukan oleh system analyst dan programmer dalam membangun sebuah sistem informasi (Beno et al., 2022). Menurut para ahli, model watelfall adalah pendekatan langkah demi langkah untuk menganalisis dan membuat desain sistem menggunakan siklus khusus untuk aktivitas pengguna. Berdasarkan penjelasan di atas, waterfall dapat menyimpulkan bahwa Sebuah siklus di mana persyaratan dan proses pengguna dipahami dan dipilih, dan sistem dibangun pada setiap tahap perencanaan, analisis, desain, dan implementasi sehingga dapat digunakan untuk memenuhi kebutuhan pengguna.



Gambar 1. Bagan Waterfall

2.2 Perancangan Sistem

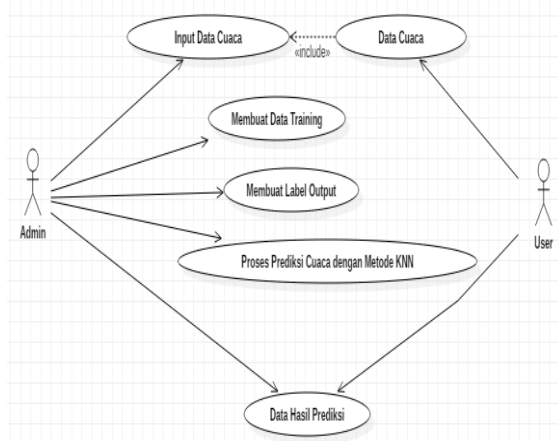
Langkah-langkah yang diambil dalam merancang sistem ini membuat saran logis dan lainnya untuk pemecahan masalah. Berikut ini adalah diagram Analisis Pengujian Metode Knn (K-Nearest Neighbors) Pada Studi Kasus Prediksi Cuaca Menggunakan Model Data Mining.



Gambar 2. Dokumen flowchart sistem
 Gambar diatas merupakan rancangan sistem yang diusulkan dimana alur sistem dimulai dari admin menginput data cuaca yang terdiri dari data temperatur, data kelembaban, data kecepatan angin. Setelah menginput data cuaca selanjutnya dilakukan proses prediksi dengan pembuatan data testing. Hasil prediksi dapat diberikan kepada masyarakat sebagai informasi perkiraan cuaca.

2.3 Use Case Sistem

Upayah-upayah yang diambil dalam merancang sistem ini membuat saran logis dan lainnya untuk pemecahan masalah. Berikut ini adalah blok use case Analisis Metode Knn (K–Nearest Neighbors) Pada Studi Kasus Prediksi Cuaca Menggunakan Model Data Mining.

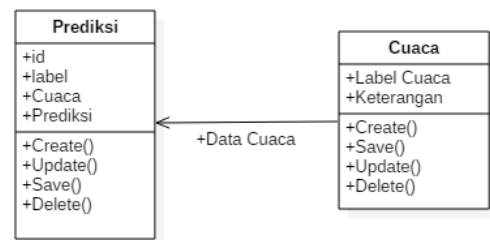


Gambar 3. Use Case Sistem
 Pada gambar diatas merupakan sistem yang diusulkan, dimana admin melakukan

penginputan data cuaca. Selanjutnya admin akan memasukkan data cuaca ke dalam file excel kemudian data cuaca akan dilakukan proses data training dengan label dan kategori. Setelah itu dilakukan pre-processing dilanjutkan proses klasifikasi prediksi Cuaca dengan metode KNN. Selanjutnya menampilkan hasil Prediksi Cuaca.

2.4 Class Diagram

Class diagram menggambarkan struktur dan deskripsi class, package dan objek beserta hubungan satu sama lain seperti containment, pewarisan, asosiasi, dan lain-lain. Dibawah ini merupakan class diagram Analisis Metode Knn (K–Nearest Neighbors) Pada Studi Kasus Prediksi Cuaca Menggunakan Model Data Mining :



Gambar 4. Class Diagram (Ismail, 2020)
 Pada gambar diatas menunjukkan class diagram sistem. Pada class diagram terdapat 2 kelas. Class prediksi dan class cuaca

3. HASIL DAN PEMBAHASAN

Setelah dilakukan perancangan sistem. Selanjutnya dilakukan implementasi sistem untuk mendapatkan hasil. Penelitian ini bertujuan menghasilkan sistem Implementasi Metode K–Nearest Neighbors Untuk Prediksi Cuaca dengan Konsep Data Mining. Untuk mencapai hasil sistem yang bagus dilakukan beberapa tahapan implementasi:

3.1 Pembuatan Dataset

Dataset adalah elemen yang penting dari seluruh proses data analytics. Dataset yang terstruktur dengan baik akan sangat membantu hasil data analysis yang dilakukan, membuat jadi lebih cepat, efisien, dan akurat. Berikut dataset Implementasi Metode K–Nearest Neighbors Untuk Prediksi Cuaca dengan Konsep Data Mining:

Tabel 1: Dataset Data Cuaca

Temperatur	Kecepatan	Kelembaban	Keterangan
20	10	26	?
24	12	25	?
22	10	26	?
21	11	27	?
25	12	27	?
18	13	28	?
18	14	22	?

Tabel diatas merupakan dataset data cuaca bulan maret tahun 2024. Jumlah dataset yang digunakan pada penelitian Penerapan Data Mining untuk prediksi cuaca Menggunakan Metode KNN sebanyak 8 data. Dataset merupakan data awal yang diproses pada sistem penerapan data mining. Fungsi dataset untuk mendapatkan model awal dalam prediksi. Pada dataset diatas akan dilanjutkan ketahap berikutnya yaitu modeling data dengan membuat label output. Untuk keterangan akan dilakukan prediksi dengan output **Hujan** atau **Tidak**.

3.2 Pembuatan Data Testing

Pembuatan data testing merupakan proses pengujian anantara output data training dengan hasil data testing. Data testing dilakukan dengan menggunakan modeling KNN. Berikut hasil data testing:

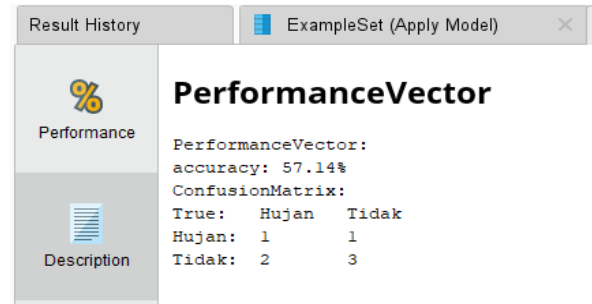
Row No.	Keterangan	prediction(Keterangan)	confidence(Hujan)	confidence(Tidak)	Temperatur	Kecepatan	Kelembaban
1	Hujan	Hujan	0.575	0.425	20	10	26
2	Hujan	Tidak	0.416	0.584	24	12	25
3	Tidak	Tidak	0.381	0.619	22	10	26
4	Tidak	Hujan	0.540	0.460	21	11	27
5	Tidak	Tidak	0.378	0.624	25	12	27
6	Hujan	Tidak	0.447	0.553	18	13	28
7	Tidak	Tidak	0.382	0.618	18	14	22

Gambar 5. Data Testing

Gambar diatas merupakan proses pembuatan data testing menggunakan algoritma KNN. Dari hasil yang dilakukan, anantara data training dan data testing 33% menghasilkan yang sama dan menghasilkan confidence (Y) dan 75% confidence (T). hasil prediksi dapat dilihat dalam bentuk grafik dan tabel. Tingkat akurasi 57%.

3.3 Hasil Prediksi

Hasil prediksi cuaca dapat dilihat dalam bentuk grafik maupun tabel. Dari hasil prediksi terdapat prediksi hujan, dan tidak hujan. Berikut hasil Analisis Prediksi Cuaca Menggunakan Model Data Mining:



Gambar 5. Hasil Prediksi

Gambar diatas merupakan hasil Deskripsi prediksi cuaca di kantor dinas pertanian kabupaten Soppeng. Hasil yang didapatkan sebanyak 33% prediksi hujau dan 75% prediksi tidak. Tingkat akurasi 57.14%.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan tentang Implementasi Metode Knn (K-Nearest Neighbors) Prediksi Cuaca Menggunakan Model Data Mining, maka dapat ditarik kesimpulan : Hasil perancangan sistem dengan menggunakan *Data Mining* dapat memudahkan proses prediksi cuaca. Perancangan sistem dapat membentuk kelompok data yang konsisten sehingga memudahkan untuk diolah lebih lanjut dalam prediksi menggunakan algoritma KNN. Hasil implementasi Sistem Analisis Pengujian Metode Knn (K-Nearest Neighbors) Pada Studi Kasus Prediksi Cuaca Menggunakan Model Data Mining berjalan dengan baik dan menghasilkan tiga bentuk prediksi berdasarkan prediksi yaitu hujan dan tidak hujan. Hasil implementasi mendapatkan 33% Prediksi hujan, dan 75% Prediksi tidak hujan dari total data cuaca sebanyak 8 Data.

5. DAFTAR PUSTAKA

- Agung, A. S., Fauzi, A. A., Nur Risal, A. A., & Adiba, F. (2023). Implementasi Teknik Data Mining terhadap Klasifikasi Data Prediksi Curah Hujan BMKG Di Sulawesi Selatan. *Jurnal Tekno Insentif*, 17(1), 22–23. <https://doi.org/10.36787/jti.v17i1.955>
- Beno, J., Silen, A. ., & Yanti, M. (2022).

- PERBANDINGAN MODEL WATERFALL DENGAN PROTOTYPE PADA PENGEMBANGAN SYSTEM INFORMASI BERBASIS WEBSITE. *JURNAL FUSION*, 33(1), 1–12.
- Dwiyanti, Z. A., & Prianto, C. (2023). Prediksi Cuaca Kota Jakarta Menggunakan Metode Random Forest. *Jurnal Tekno Insentif*, 17(2), 127–137. <https://doi.org/10.36787/jti.v17i2.1136>
- Gelinas, Ulric, Oram, Alan, Wiggins, & William. (2020). PERBANDINGAN ALGORITME KLASIFIKASI UNTUK PREDIKSI CUACA. *Accounting Information System*, 3(1), 17–30.
- Ismail. (2020). Penerapan Metode Rapid Application Development dalam Perancangan Sistem Informasi Pendataan. *Jurnal Infokam*, 16(1), 46–55. <http://amikjtc.com/jurnal/index.php/jurnal/article/view/218/164#>
- Ismail Ismail, Rezky Erwin Syah, M. A. T. (2024). Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI)*, 7(2), 270–281.
- Perdana, D. S. (2022). PENERAPAN METODE K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN KETAHANAN PANGAN. *Jurnal Sistem Informasi (JUSIN)*, 3(2), 67–72.
- Purnamawati, A., Winnarto, M. N., & Mailasari, M. (2022). Analisis Cart (Classification and Regression Trees) Untuk Prediksi Pengguna Sepeda Berdasarkan Cuaca. *Jurnal Teknoinfo*, 16(1), 14. <https://doi.org/10.33365/jti.v16i1.1478>
- Rofiq, H., Pelangi, K. C., & Lasena, Y. (2020). Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma Naive Bayes. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 3(1), 8–15. <http://mahasiswa.dinus.ac.id/docs/skripsi/jurnal/19417.pdf>
- Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naive Bayes : Systematic Review. *Faktor Exacta*, 13(1), 35. <https://doi.org/10.30998/faktorexacta.v13i1.5548>
- Tita Lattifia, Putu Wira Buana, & NI Kadek Dwi Rusjayanthi. (2022). Model Prediksi Cuaca Menggunakan Metode LSTM. *JITTER- Jurnal Ilmiah Teknologi Dan Komputer*, 3(1).
- Udjulawa, D. (2023). Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Cuaca. *Jurnal Algoritme*, 4(1), 1–12. <https://doi.org/10.35957/algoritme.xxxx>
- Ulinnuha, N., & Farida, Y. (2018). Prediksi Cuaca Kota Surabaya Menggunakan Autoregressive Integrated Moving Average (Arima) Box Jenkins dan Kalman Filter. *Jurnal Matematika "MANTIK,"* 4(1), 59–67. <https://doi.org/10.15642/mantik.2018.4.1.59-67>