

Klasifikasi Cuaca di DKI Jakarta Menggunakan Metode Gaussian Naive Bayes

Andhika Fariz Irwanto¹, Audi Raditama², Yoga Fadillah Putra Majid³

Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya

Email : andhikafariz16@gmail.com, audhie240702@gmail.com, yogafdllh00@gmail.com

ABSTRAKSI

Klasifikasi Cuaca dapat membantu masyarakat dan pihak terkait dalam mengambil keputusan yang tepat, terutama dalam perencanaan aktivitas sehari-hari. Metode Naïve bayes ini digunakan dalam penelitian ini karena keunggulannya dalam klasifikasi pada data yang memiliki banyak atribut. Data cuaca yang digunakan dalam penelitian ini meliputi variabel seperti kelembapan, suhu derajat celsius, cuaca. Model naïve bayes diuji menggunakan data pengujian keakuratan model dievaluasi dengan menghitung metric evaluasi akurasi. Oleh karena itu penentuan cuaca untuk mendapatkan informasi cuaca perlu dibuat sehingga dapat dimanfaatkan oleh masyarakat. Hasil penelitian menunjukkan bahwa Metode Gaussian Naïve Baiyes dapat mengklasifikasikan kondisi cuaca DKI Jakarta dengan akurasi sebesar 81,25%. Hasil Evaluasi ini memberikan gambaran tentang seberapa baik model naïve bayes dalam melakukan klasifikasi cuaca DKI Jakarta.

Kata Kunci: Cuaca, Klasifikasi, Naive Bayes, Gaussian, Dataset.

ABSTRACT

Weather Classification can help the community and related parties in making the right decisions, especially in planning daily activities. The Naïve Bayes method is used in this study because of its superiority in classifying data that has many attributes. The weather data used in this study includes variables such as humidity, degree Celsius temperature, weather. The naïve Bayes model is tested using test data for the accuracy of the model and evaluated by calculating the accuracy evaluation metric. Therefore, weather determination to obtain weather information needs to be made so that it can be utilized by the community. The results showed that the Gaussian Naïve Baiyes method could classify DKI Jakarta weather conditions with an accuracy of 81.25%. The results of this evaluation provide an overview of how good the Naïve Bayes model is in classifying DKI Jakarta weather.

Keywords: Weather, Classification, Naive Bayes, Gaussian, Dataset.

Penulis Korespondensi

Audi Raditama

This is an open access article under the [CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license



Tanggal Submit : 14/07/2023
Tanggal Diterima : 30/08/2023
Tanggal Terbit : 31/08/2023

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) International License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).

Publisher's Note: JPPM stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

I. PENDAHULUAN

Cuaca menjadi faktor penting dalam segala aspek kehidupan. Perubahan cuaca yang signifikan berpengaruh terhadap kegiatan manusia. Prediksi cuaca yang akurat sangat penting untuk berbagai kepentingan, seperti keamanan penerbangan, pengaturan lalu lintas, persiapan bencana, dan kegiatan luar ruangan lainnya[1]. Perubahan cuaca bergantung pada banyak faktor seperti suhu, kelembaban, angin, waktu, lokasi

dan lainnya. Dari faktor ini dapat diperoleh jenis cuaca seperti cuaca +cerah, berawan, hujan, hujan petir. Cuaca ekstrim berpotensi menjadi bencana seperti banjir, tanah longsor, kebakaran, penyebaran penyakit yang mempengaruhi daya tahan tubuh manusia. Dengan teknik klasifikasi cuaca yang baik dapat diprediksi kemungkinan perubahan cuaca yang terjadi dengan lebih akurat.

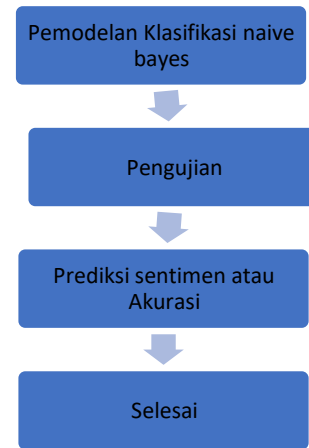
Dalam prediksi cuaca DKI Jakarta, metode naïve bayes digunakan untuk mengklasifikasikan cuaca di wilayah DKI Jakarta dan memperhitungkan data cuaca selama periode waktu tertentu agar menghasilkan hasil yang akurat dan efisien. Tujuan klasifikasi cuaca di DKI Jakarta menggunakan metode Naive Bayes adalah untuk menghasilkan model klasifikasi yang akurat dan efisien untuk memprediksi kondisi cuaca di DKI Jakarta. Dengan menggunakan model klasifikasi ini, dapat membantu berbagai aplikasi, seperti transportasi, pertanian, pariwisata, dan manajemen bencana dalam mengambil keputusan yang tepat dan mempersiapkan diri menghadapi kondisi cuaca yang mungkin terjadi. Beberapa peneliti juga menggunakan metode naïve bayes untuk mengklasifikasikan prakiraan curah hujan [2], mengklasifikasikan komentar spam pada Youtube [3], mengklasifikasikan jenis jamur [4], dan mengetahui kualitas air di DKI Jakarta [5]. Beberapa penelitian yang membahas tentang cuaca juga banyak yang menggunakan metode lain seperti KNN[6], SVM dan NBC[7], serta menggunakan metode CNN[8].

II. PENELITIAN YANG TERKAIT

Dari penelitian yang dilakukan oleh I Gede Aris Gunadi, Ayu Aprilyana Kusuma Dewi Metode Naïve Bayes bekerja dengan menghitung peluang dari satu kelas dari masing-masing kelompok atribut yang ada dan menentukan kelas mana yang paling optimal, artinya pengelompokan dapat dilakukan berdasarkan kategori ringan, normal, dan deras berturut turut sebesar 79.5%, 40.9%, dan 86.4%[9]. Untuk mengetahui cuaca juga dibutuhkan parameter seperti suhu, kecepatan angin, dan curah hujan yang menjadi komponen utama untuk cuaca[10].

III. METODE PENELITIAN

Naive bayes merupakan metode pengklasifikasian berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel independen. Pada algoritma ini pembelajaran lebih ditekankan pada pengestimasi probabilitas. Keuntungan algoritma *naive bayes* adalah tingkat nilai error yang didapat lebih rendah ketika dataset berjumlah besar, selain itu akurasi *naive bayes* dan kecepatannya lebih tinggi pada diaplikasikan ke dalam dataset yang jumlahnya lebih besar.



Gambar 1. Tahapan Penelitian

Distribusi Gaussian adalah asumsi pendistribusian nilai kontinu yang terkait dengan setiap fitur berisi nilai numerik. Ketika diplot, akan muncul kurva berbentuk lonceng yang simetris tentang rata-rata nilai fitur.

IV. HASIL DAN PEMBAHASAN

4.1. Analisis System

Analisa system memegang peranan penting dalam konstruksi dan desain system. Analisis sistematis adalah jalan menuju pengetahuan bagaimana dan masalah apa yang ada di system sehingga akan ditemukan solusinya setelah wawancara dan kekurangan dicatat dalam pengamatan penulis dalam ssistem perkiraan cuaca BMKG, Sistem yang mana berjalan di BMKG, tidak ada proses Otomatisasi lengkap operasi Ramalan. BMKG membuat prakiraan cuaca dengan memperhatikan kondisi cuaca yang ada terjadi, misalnya, dalam kondisi suhu kelembaban dan kecepatan angin. Setelah Pemantauan kondisi

cuaca yang dianalisa oleh BMKG diperoleh kondisi cuaca sebelumnya Dapatkan hasil ramalan cuaca. Kemudian Cek di website BMKG. Berhubungan dengan Sistem kemudian dibangun dari analisis di atas yang dapat melakukan proses peramalan cuaca secara terus menerus secara otomatis menggunakan metode Naive Bayes klasifikasi.

4.2 Analisis Proses Perhitungan

Dapatkan analisis proses perhitungan nilainya harus diuji dengan membandingkan hasilnya uji sistem manual ya Perbandingan ini menghasilkan hasil yang sistematis yaitu setelah perhitungan manual sepenuhnya dan dapat dibandingkan dengan hasil perhitungan Aplikasi.

1. Perhitungan naïve bayes dengan Gaussian
2. Mendapatkan hasil accuracy

Sebagai contoh analisis yang akan dilakukan perhitungan manual, diambil dari contoh dataset yang dapat dilihat di Table 1.

Tabel 1. Dataset Penelitian

| tanggal | wilayah | waktu | cuaca | kelembaban_perse | suhu_d erajat_celcius |
|----------|------------------|-----------|---------------|------------------|-----------------------|
| 8/3/2018 | Jakarta Pusat | Dini Hari | Berawan | 40 – 95 | 22 – 33 |
| 8/3/2018 | Jakarta Utara | Dini Hari | Berawan | 40 – 95 | 22 – 33 |
| 8/4/2018 | Jakarta Timur | Pagi | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Barat | Pagi | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Selatan | Pagi | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Kepulauan Seribu | Pagi | Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Pusat | Pagi | Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Utara | Pagi | Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Timur | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Barat | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Selatan | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Kepulauan Seribu | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Pusat | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Utara | Siang | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Timur | Malam | Cerah Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Barat | Malam | Berawan | 45 – 80 | 22 – 33 |
| 8/4/2018 | Jakarta Selatan | Malam | Cerah Berawan | 45 – 80 | 22 – 33 |

```

confusion matrix:
[[10 0 0 0 0 0 0 0 0 0 0]
 [ 0 13 0 0 0 0 0 6 0 0 0]
 [ 0 0 5 0 0 0 0 0 0 0 0]
 [ 0 0 0 15 0 0 0 0 0 0 0]
 [ 0 0 0 0 5 0 0 0 0 0 0]
 [ 0 0 0 0 0 5 0 0 0 0 0]
 [ 0 0 0 0 0 0 22 0 0 0 0]
 [ 0 4 0 0 0 0 0 19 0 0 0]
 [ 0 0 0 0 0 0 0 0 9 0 0]
 [ 0 0 0 0 0 0 0 0 0 9 0]
 [ 0 0 0 0 0 0 0 0 0 5]]

```

```

Classification report:
              precision    recall  f1-score   support

22 □ 32         1.00        1.00        1.00         10
22 □ 33         0.62        0.68        0.65         19
22 □ 34         1.00        1.00        1.00          5
23 □ 32         0.54        0.79        0.64         19
23 □ 33         1.00        1.00        1.00          5
23 □ 34         1.00        1.00        1.00          5
24 □ 32         0.79        0.76        0.77         29
24 □ 33         1.00        0.66        0.79         29
24 □ 34         1.00        1.00        1.00          9
25 □ 33         1.00        1.00        1.00          9
25 □ 34         1.00        1.00        1.00          5

accuracy              0.90
macro avg             0.90
weighted avg          0.85

```

Accuracy Score: 81.250000%

Histogram terbagi menjadi beberapa interval atau kelas yang saling berdekatan, dan tinggi setiap batang pada histogram menunjukkan frekuensi atau jumlah data yang jatuh dalam interval tersebut. Semakin tinggi batang histogram, semakin banyak data yang termasuk dalam interval tersebut. Histogram digunakan untuk menganalisis distribusi data dan memberikan gambaran tentang bagaimana data tersebar. Dengan melihat histogram, kita dapat mengidentifikasi apakah data

memiliki distribusi normal, asimetri, atau ada anomali dalam data tersebut.

V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa dengan metode Gaussian Naïve Bayes kita bisa melakukan Klasifikasi cuaca di DKI Jakarta, dan mendapatkan tingkat akurasi sebesar 81,25%

DAFTAR PUSTAKA

- [1] F. Hamami and A. Dahlan, "KLASIFIKASI CUACA PROVINSI DKI JAKARTA MENGGUNAKAN ALGORITMA RANDOM FOREST DENGAN TEKNIK OVERSAMPLING," 2022.
- [2] "HALAMAN SAMPUL PENERAPAN ALGORITMA NAÏVE BAYES CLASSIFIER DALAM." [Online]. Available: <http://digilib.mercubuana.ac.id/>
- [3] Y. Musa and dan Yaya Wihardi, "KLASIFIKASI KOMENTAR SPAM PADA YOUTUBE MENGGUNAKAN METODE NAÏVE BAYES, SUPPORT VECTOR MACHINE, DAN K-NEAREST NEIGHBORS," 2108.
- [4] A. A. Mahran, R. K. Hapsari, and H. Nugroho, "PENERAPAN NAIVE BAYES GAUSSIAN PADA KLASIFIKASI JENIS JAMUR BERDASARKAN CIRI STATISTIK ORDE PERTAMA."
- [5] Y. S. Sari, "Penerapan Metode Naïve Bayes Untuk Mengetahui Kualitas Air Di Jakarta," *Jurnal Ilmiah FIFO*, vol. 13, no. 2, p. 222, Nov. 2021, doi: 10.22441/fifo.2021.v13i2.010.
- [6] A. S. Utami, D. P. Rini, and E. Lestari, "Prediksi Cuaca di Kota Palembang Berbasis Supervised Learning Menggunakan Algoritma K-Nearest Neighbour."
- [7] "PERBANDINGAN HASIL KLASIFIKASI CURAH HUJAN MENGGUNAKAN METODE SVM DAN NBC".
- [8] R. Sulisty Budi, R. Patmasari, and S. Saidah, "KLASIFIKASI CUACA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN) WEATHER CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK (CNN) METHOD."
- [9] "Klasifikasi Curah Hujan di Provinsi Bali Berdasarkan Metode Naïve Bayesian".
- [10] A. M. Siregar, "Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning," *PETIR*, vol. 13, no. 2, pp. 138–147, Sep. 2020, doi: 10.33322/petir.v13i2.998.