

Klasifikasi Curah Hujan Iklim Mikro Dengan *Naive Bayes* Pada Wilayah Kebumen

¹Anandita Ayu Rochani*, ²Evan Alif Widhyatma, ³Rahmat Hidayat

^{1,2,3} Prodi Sains Data , Sains dan Teknologi, Universitas Putra Bangsa

*e-mail: ayat150190@gmail.com

Abstrak

Perubahan pola curah hujan yang tidak menentu sangat berpengaruh terhadap berbagai aspek kehidupan dimana segala aktivitas penting berada di dalamnya, sehingga perlu dilakukan prediksi curah hujan agar tidak mengganggu aktivitas penting dan harus segera dilaksanakan. Segala sesuatu yang berlebihan tentunya berdampak buruk, sama halnya dengan kondisi cuaca. Hujan yang berkepanjangan dapat menyebabkan banjir dan tanah longsor sedangkan kekeringan yang ekstrem dapat menyebabkan gagal panen dan krisis pangan. Karena hal tersebut maka diperlukan suatu sistem terpadu yang dapat menunjang kegiatan manusia seperti misalnya dalam pertanian, sebagai sumber kebutuhan pokok kehidupan manusia, pertanian memegang peran penting sehingga kelangsungan dalam pertanian harus terjaga untuk mencegah krisis pangan di masa mendatang.

Kata kunci: Curah Hujan, Cuaca, Pembelajaran Mesin, Naive Bayes

Abstract

Erratic changes in rainfall patterns greatly influence various aspects of life where all important activities are involved, so it is necessary to predict rainfall so that it does not interfere with important activities and must be implemented immediately. Anything in excess certainly has a bad impact, the same as weather conditions. Prolonged rains can cause floods and landslides while extreme droughts can cause crop failures and food crises. Because of this, an integrated system is needed that can support human activities, such as in agriculture, as a source of basic needs for human life, agriculture plays an important role so that continuity in agriculture must be maintained to prevent a food crisis in the future.

Keywords: Rain, Weather, Climate, Agriculture, Predictions

1 Pendahuluan

Keadaan cuaca menjadi daya tarik bagi para peneliti dalam beberapa tahun terakhir. Oleh karena itu penting untuk mengetahui cuaca dalam beberapa hari mendatang untuk mengambil tindakan pencegahan. Peramalan cuaca di masa depan khususnya curah hujan mendapat perhatian banyak peneliti, guna mencegah banjir dan risiko lain yang timbul dari curah hujan [1]. Kondisi cuaca sangat berpengaruh dalam kehidupan sehari-hari, seperti dalam bidang pertanian, transportasi dan industri. Dalam bidang pertanian, faktor cuaca dijadikan sebagai salah satu pertimbangan dalam penentuan kecocokan jenis tanaman yang akan dibudidayakan. Memprediksi kejadian hujan sangat penting, namun memprediksi curah hujan merupakan masalah yang kompleks karena dipengaruhi oleh sifat dinamis parameter cuaca tropis yaitu tekanan atmosfer, suhu, kelembaban, titik embun dan kecepatan angin [2]. Maka dari itu pengamatan terhadap kondisi cuaca, khususnya kondisi curah hujan sangat penting dilakukan. Kajian prakiraan cuaca merupakan topik penelitian yang menarik karena cuaca mempengaruhi kehidupan sehari-hari dan keakuratan prakiraan cuaca sangat penting. Prakiraan cuaca dilakukan dengan mengumpulkan data kuantitatif tentang keadaan atmosfer saat ini dan menggunakan pengetahuan ilmiah tentang proses atmosfer untuk memprediksi bagaimana atmosfer akan berevolusi.

Kabupaten Kebumen merupakan salah satu wilayah pesisir Barat Daya Jawa Tengah yang memiliki karakteristik curah hujan yang tinggi, rata-rata curah hujannya yaitu 3.250 mm per tahun [3]. Curah hujan yang tinggi memberikan dampak besar terhadap kehidupan, seperti penerbangan, perkebunan, dan pertanian. Curah hujan yang di atas rata-rata dapat menghambat hasil pertanian dan perkebunan, sedangkan jika curah hujan berada di bawah rata-rata maka dapat menghambat pertumbuhan tanaman akibat dari kekurangan air. Petani membutuhkan informasi mengenai curah hujan untuk mencari alternatif jika terjadi masalah curah hujan seperti kekeringan dan kelebihan curah

hujan. Dalam upaya ini peningkatan ketangguhan sistem peringatan dini cuaca sangat diperlukan, penelitian mengenai prediksi probabilitas hujan menjadi semakin penting, terutama di wilayah geografis yang rentan terhadap perubahan cuaca, seperti Kabupaten Kebumen.

Dalam rangka meningkatkan akurasi prediksi probabilitas hujan, teknik-teknik pemodelan seperti Naive Bayes telah menjadi fokus utama penelitian. Naive Bayes Klasifier cocok untuk meramalkan curah hujan karena tidak memerlukan pemenuhan asumsi seperti metode klasifikasi lainnya seperti analisis diskontinuitas. Penelitian terkait sebelumnya telah dilakukan oleh dengan judul " *Naive bayes methods for rainfall prediction classification in Banyuwangi*". Hasil penelitian ini adalah klasifikasi prediksi curah hujan pada kategori hujan ringan, normal, dan lebat, sistem mampu mengklasifikasikan dengan baik [4]. Kemungkinan masih rendahnya rendemen untuk kategori hujan normal terletak pada data, seharusnya tidak ada perbedaan yang signifikan antara ketiga kelas. Perlu dilakukan pengujian ulang dengan lebih banyak data.

Dibandingkan dengan metode konvensional memprediksi curah hujan dengan Artificial Intelligent (AI) akan membantu dalam prediksi cepat dan akurat. Namun, fenomena curah hujan bukanlah fenomena yang dapat dipelajari secara sederhana, hal ini disebabkan karena banyak faktor yang mempengaruhi curah hujan di suatu daerah [5]. Prediksi adalah subjek yang menarik untuk penelitian meskipun iklim adalah salah satu faktor yang mempengaruhi kehidupan sehari-hari manusia, sehingga ketepatan dalam prakiraan sangat penting. Laporan cuaca dihasilkan dengan mengumpulkan informasi empiris tentang kondisi meteorologi saat ini dan mencoba memproyeksikan seberapa baik atmosfer akan tumbuh dengan menggunakan pengetahuan ilmiah tentang fisika atmosfer. Penelitian peramalan telah dilakukan beberapa kali dengan menggunakan beberapa teknik seperti logika fuzzy Sugeno dan Mamdani, Naive Bayes, dan C.45 [6].

Berdasarkan poin-poin di atas penulis tertarik untuk melakukan penelitian mengenai klasifikasi prediksi curah hujan di Kebumen, karena pembahasan mengenai masalah prediksi curah hujan di Kebumen masih sedikit. Selanjutnya, masalah dievaluasi untuk menemukan solusi. Maka peneliti ingin melakukan penelitian ini, peneliti juga mempelajari dasar teori dari berbagai literatur mengenai penerapan metode Naive Bayes, ide dan teori data mining, dan prediksi curah hujan di Kebumen. Selain itu, peneliti menggunakan dasar jurnal-jurnal terdahulu yang membahas mengenai masalah prediksi curah hujan.

2 Tinjauan Literatur

Perkembangan teknologi informasi telah memungkinkan pemanfaatan teknik-teknik komputasi untuk memprediksi berbagai fenomena alam, termasuk curah hujan. Prediksi curah hujan memiliki peran krusial dalam berbagai sektor, mulai dari pertanian, mitigasi bencana, hingga manajemen sumber daya air, terutama pada skala iklim mikro yang memerlukan akurasi tinggi. Pemanfaatan *machine learning* (pembelajaran mesin) menjadi pendekatan yang semakin populer dalam mengatasi kompleksitas data meteorologi. *Machine learning* memberikan kemampuan kepada sistem komputer untuk "belajar" dari data, memodifikasi program secara otomatis agar menghasilkan *output* yang lebih akurat, tanpa memerlukan pemrograman eksplisit [7]. Hal ini memungkinkan sistem untuk menghasilkan pengetahuan baru dari data yang dikumpulkan dan memecahkan masalah yang sulit ditentukan dengan aturan konvensional [8]. Penerapan *machine learning* untuk prediksi cuaca, khususnya curah hujan, seringkali melibatkan tahapan *data mining* atau penemuan pengetahuan dalam basis data (KDD). *Data mining* adalah proses ekstraksi dan identifikasi informasi serta pola yang berguna dari basis data besar menggunakan teknik statistik, kecerdasan buatan, dan *machine learning*.

Berbagai algoritma *machine learning* telah dieksplorasi untuk klasifikasi dan prediksi curah hujan. Salah satu algoritma yang menunjukkan kinerja baik dalam tugas klasifikasi berbasis probabilitas adalah Naive Bayes [9]. Naive Bayes bekerja dengan mengestimasi himpunan peluang berdasarkan kemunculan dan kombinasi nilai dalam himpunan data, dengan asumsi bahwa semua atribut bersifat independen. Keunggulan Naive Bayes terletak pada efisiensinya dalam hal ruang dan waktu komputasi, kemampuannya menangani atribut yang tidak relevan, serta tidak memerlukan data dalam jumlah sangat besar untuk estimasi parameter. Selain itu, algoritma ini cepat dalam memprediksi kelas berdasarkan input dan sangat cocok untuk variabel kategorik atau nominal, bahkan pada dataset yang relatif kecil.

Meskipun banyak penelitian telah menerapkan *machine learning* untuk prediksi curah hujan pada skala regional atau nasional, studi mengenai klasifikasi curah hujan iklim mikro, khususnya di wilayah

spesifik seperti Kebumen, masih terbatas. Penelitian sebelumnya seringkali berfokus pada prediksi umum atau wilayah dengan karakteristik geografis yang berbeda, sehingga akurasi dan relevansinya untuk iklim mikro lokal menjadi pertanyaan. Belum banyak penelitian yang secara spesifik menguji kinerja algoritma Naive Bayes untuk klasifikasi curah hujan pada data iklim mikro di Kebumen, yang memiliki karakteristik geografis dan topografi unik yang dapat mempengaruhi pola curah hujan. Oleh karena itu, artikel ini bertujuan untuk mengisi kesenjangan tersebut dengan mengembangkan model klasifikasi curah hujan iklim mikro menggunakan algoritma Naive Bayes pada wilayah Kebumen. Penelitian ini akan berfokus pada bagaimana Naive Bayes dapat secara efektif mengklasifikasikan pola curah hujan pada skala yang lebih detail, mempertimbangkan karakteristik data iklim mikro lokal, dan memberikan kontribusi pada pemahaman serta aplikasi *machine learning* untuk manajemen sumber daya air dan mitigasi bencana di tingkat lokal.

3 Metode Penelitian

Penelitian yang dilakukan menggunakan data curah hujan di Kebumen dari tahun 2010 hingga 2023 dari Stasiun Meteorologi Tunggul Wulung Cilacap milik BMKG (Badan Meteorologi Klimatologi dan Geofisika). Penelitian ini dimulai dengan pengamatan dan identifikasi masalah. Analisis dilakukan setelah data terkumpul untuk menemukan pendekatan yang paling sesuai dan membantu proses menemukan solusi yang sesuai dengan tujuan peneliti. Data yang didapatkan akan diolah dengan perhitungan manual dibantu dengan Microsoft Excel kemudian dilakukan pengujian dibantu menggunakan perangkat lunak Orange3. Setelah melewati proses pengolahan data dan pengujian maka dapat ditarik kesimpulan akan hasil dari proses penelitian yang dilakukan. Berikut merupakan alur kerja penelitian yang dilakukan.

Studi literatur dan Mengidentifikasi Masalah

Dalam hal ini kami melihat masalah yang terkait dengan klasifikasi prediksi curah hujan di Kebumen karena pembahasan mengenai masalah prediksi curah hujan di Kebumen masih sedikit. Selanjutnya, masalah dievaluasi untuk menemukan solusi. Maka peneliti ingin melakukan penelitian ini, peneliti juga mempelajari dasar teori dari berbagai literatur mengenai penerapan metode Naive Bayes, ide dan teori data mining, dan prediksi curah hujan di Kebumen. Selain itu, peneliti menggunakan dasar jurnal-jurnal terdahulu yang membahas mengenai masalah prediksi curah hujan.

Pengumpulan Data

Pada penelitian ini melalui menggunakan metode penelitian kuantitatif. Data yang digunakan dalam penelitian ini adalah data curah hujan di Kebumen dari tahun 2010 hingga 2023 dari Stasiun Meteorologi Tunggul Wulung Cilacap milik BMKG (Badan Meteorologi Klimatologi dan Geofisika) melalui website [visualcrossing](https://visualcrossing.com/). Penelitian kuantitatif ini berfokus pada penggunaan angka, tabel, grafik, dan diagram untuk menunjukkan hasil data yang dikumpulkan dan digunakan sebagai bahan analisis data metode Naive Bayes.

Klasifikasi Naive Bayes

Klasifikasi juga merupakan pengelompokan data dimana data yang digunakan mempunyai kelas atau target yang akan diklasifikasi. Sehingga nantinya algoritma dapat digunakan untuk menyelesaikan masalah yang akan klasifikasi [10]. Algoritma Naive Bayes [11] itu sendiri merupakan salah satu algoritma yang digunakan untuk teknik klasifikasi data. Naive Bayes digunakan untuk teknik pengklasifikasian data dengan metode kemungkinan dan statistik yang memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive Bayes dimana diasumsikan kondisi antara atribut yang satu dengan atribut yang lainnya itu saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak adanya ciri-ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas yang lainnya [10]. Tahapan dalam metode Naive Bayes :

- 1) Pengecekan data cuaca apakah data set adalah data yang konsisten, bias, terdapat null atau tidak konsisten
- 2) Setelah dilakukan analisa terhadap data cuaca, maka dapat disimpulkan bahwa data set adalah data yang tidak konsisten maupun bias sehingga kita tidak dapat membuat hipotesis langsung.

- 3) Metode Bayes menggunakan probabilitas bersyarat sebagai dasarnya. Probabilitas bersyarat dinyatakan sebagai :

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Keterangan:

X, Y = Peristiwa

$P(Y|X)$ = Probabilitas Y berdasarkan kondisi X

$P(X|Y)$ = Probabilitas X berdasarkan kondisi Y

$P(X)$ = Probabilitas X

$P(Y)$ = Probabilitas Y

Untuk menggambarkan teorema Naive Bayes, kita harus tahu bahwa proses klasifikasi membutuhkan berbagai arahan untuk menentukan kelas yang sesuai dengan sampel yang dianalisis.

Tahapan Implementasi

Pada tahap implementasi penelitian ini, pengolahan data diterapkan dalam sebuah tools dengan menggunakan perangkat lunak Orange3. Selanjutnya, peneliti melakukan pengujian untuk mengetahui apakah penelitian telah mencapai tujuan untuk dapat memprediksi curah hujan di Kebumen. Hasil dari berbagai tahapan sebelumnya dapat digunakan untuk membuat kesimpulan tentang hasil pengolahan data dan pengujian yang telah dilakukan. Dari hasil ini, dapat dievaluasi apakah algoritma yang digunakan sesuai untuk menemukan solusi masalah.

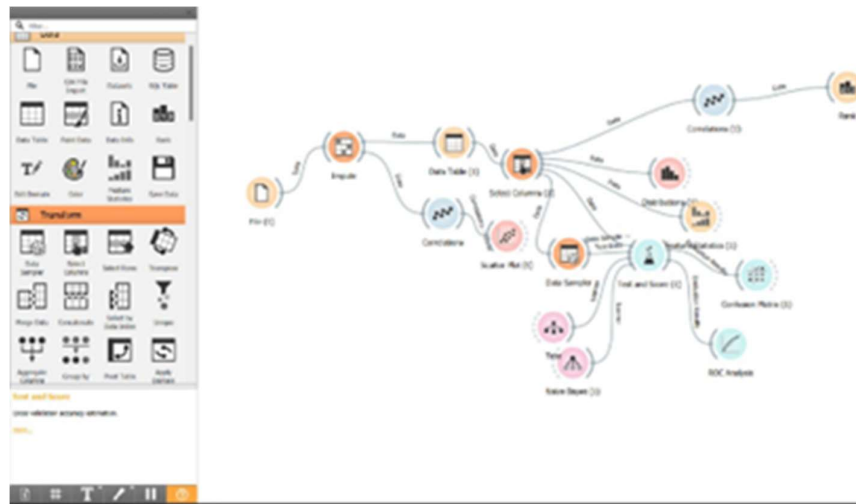
4 Hasil dan Pembahasan

Penelitian ini mengumpulkan data menggunakan algoritma naive bayes. Data ini dikumpulkan dari Stasiun Meteorologi Tunggul Wulung Cilacap milik BMKG (Badan Meteorologi Klimatologi dan Geofisika) di Kebumen dari tahun 2010 hingga 2023. Jumlah curah hujan dipengaruhi oleh sejumlah variabel, menurut data yang dikumpulkan. Dengan data ini, pengolah dan penghitungan dapat dilakukan hingga hasil perhitungan ditemukan untuk mengetahui pengelompokan data. Data penelitian ini dibersihkan secara manual dengan menggunakan perangkat lunak Microsoft Excel. Oleh karena itu, proses membersihkan data adalah tahap awal proses, yaitu menentukan dataset yang akan digunakan. Di bawah ini terdapat sampel data dari Kabupaten Kebumen yang sudah difilter yang berisi variabel yang mempengaruhi curah hujan dari tahun 2010 hingga 2023 di Kabupaten Kebumen.

	kategori_curah_hujan	temp	dew	humidity	windspeed	winddir	sealevelpressure	cloudcover	sollarradiation
1	Tidak Hujan	27.8	23.9	78.1	126	236.7	1009.5	79.0	284.0
2	Sedang	28.2	24.7	82.1	162	248.1	1010.6	82.7	163.3
3	Ringan	28.4	25.0	82.5	22	243.9	1010.3	73.6	202.1
4	Ringan	28.4	25.4	83.9	238	245.6	1009.6	72.2	144.7
5	Ringan	28.6	25.5	83.6	241	258.9	1009.1	80.7	139.0
6	Ringan	28.2	24.1	78.6	245	268.2	1010.8	71.1	203.1
7	Tidak Hujan	27.7	24.5	82.7	184	258.8	1010.4	76.1	233.3
8	Tidak Hujan	27.3	23.6	80.5	148	276.1	1010.9	77.0	239.0
9	Sedang	26.4	23.1	82.8	148	236.1	1011.3	83.2	121.4
10	Ringan	26.8	23.4	81.6	151	217.9	1011.4	74.9	185.2
11	Tidak Hujan	27.2	24.5	83.6	209	253.2	1011.7	68.6	177.7
12	Sedang	27.8	24.6	83.4	263	264.1	1011.2	76.4	178.2
13	Ringan	28.2	24.4	80.1	32	269.7	1010.9	72.1	162.6
14	Ringan	28.0	24.3	80.8	302	278.5	1012.0	57.9	273.0
15	Ringan	27.7	23.7	79.4	295	285.1	1011.7	66.2	273.4
16	Ringan	28.1	24.1	78.6	313	280.1	1011.4	65.9	276.4
17	Ringan	27.9	23.9	79.6	36	271.6	1011.0	56.4	300.4
18	Ringan	28.2	23.8	77.8	335	269.6	1010.8	73.6	195.9
19	Sedang	27.7	24.0	80.5	248	266.6	1011.5	61.3	262.7
20	Ringan	27.5	24.0	82.1	162	276.8	1012.6	68.7	207.2

Gambar 1 Dataset Curah Hujan di Kebumen

Dalam tabel berisi variabel berupa suhu , kelembapan , tekanan udara , kecepatan angin, radiasi matahari ,titik embun dan arah angin. Untuk target tujuan mengenai kategori curah hujan yang berupa tidak hujan, hujan ringan, hujan sedang, hujan lebat, hujan sangat lebat, dan hujan ekstrem.



Gambar 2 Model Naïve Bayes

Dari data tersebut peneliti melakukan pengujian model dengan metode Naïve Bayes yang menghasilkan beberapa model evaluasi yang dijabarkan di bawah ini,

Test and Score - Orange

☐ Cross validation
 Number of folds: 20
☐ Stratified
☐ Cross validation by feature
☒ Random sampling
 Repeat train/test: 20
 Training set size: 80 %
☐ Stratified
☐ Leave one out
☐ Test on train data
☐ Test on test data

Evaluation results for target (None, show average over classes)

Model	Train	Test	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
Tree	11.7...	0.012	0.642	0.583	0.584	0.585	0.583	0.279	0.693	11.915
Naive Bayes	0.368	0.066	0.785	0.653	0.643	0.638	0.653	0.385	0.726	1.202
Neural Network	92.0...	0.203	0.815	0.704	0.666	0.660	0.704	0.442	0.685	0.791

Gambar 3 Test and Score

Pada evaluasi model peneliti melakukan test score dengan EPOCH sebanyak 20 kali dan spesifikasi data training sebesar 80% disini tertera bahwa model naive bayes berada di posisi tengah-tengah antara model decision tree dan neural network. Dapat dilihat bahwa naive bayes memiliki nilai AUC sebesar 0,785 dan F1 memiliki nilai sebesar 0,643 dari nilai tersebut terlihat bahwa nilai naive bayes berada di posisi tengah antara lebih decision tree yang memiliki nilai lebih rendah dengan AUC 0,642 dan F1 dengan nilai 0,584 dan neural network dengan nilai lebih tinggi dengan AUC bernilai 0,815 dan F1 dengan nilai 0,666. Dari penjabaran tersebut kita dapat menyimpulkan prediksi curah hujan dengan EPOCH sebanyak 20 kali dan data training sebesar 80% menghasilkan prediksi seperti tabel 3.

		Predicted						
		Ekstrim	Lebat	Ringan	Sangat Lebat	Sedang	Tidak Hujan	Σ
Actual	Ekstrim	1.9 %	2.0 %	0.1 %	2.4 %	0.2 %	0.1 %	34
	Lebat	12.0 %	17.7 %	3.8 %	18.9 %	11.8 %	1.6 %	869
	Ringan	58.4 %	46.5 %	72.0 %	40.8 %	55.0 %	27.7 %	11798
	Sangat Lebat	3.2 %	7.9 %	0.5 %	8.9 %	0.9 %	0.3 %	152
	Sedang	22.4 %	24.4 %	10.0 %	28.4 %	28.6 %	2.4 %	1977
	Tidak Hujan	2.1 %	1.4 %	13.6 %	0.6 %	3.5 %	67.9 %	5630
Σ		375	492	12607	169	1140	5677	20460

Gambar 4 Confussion matrix

disimpulkan bahwa hasil prediksi curah hujan berpotensi ringan di daerah Kebumen karena memiliki presentase tertinggi bernilai 72 % dengan perolehan nilai predicted sebesar 12607 dan nilai *actual* sebesar 11798.

5 Kesimpulan

Pada penelitian ini dikembangkan model Data Mining untuk prediki curah hujan yang ada di Kabupaten Kebumen menggunakan algoritma Naive Bayes. Naïve Bayes digunakan karena kehandalannya dalam mengklasifikasi data berdasarkan atribut-atribut yang dimiliki, baik berupa nilai numerik maupun kategorik. Model dibangun menggunakan dataset curah hujan di Kebumen yang didapatkan dari Stasiun Meteorologi Tunggul Wulung Cilacap milik BMKG (Badan Meteorologi Klimatologi dan Geofisika) dan diimplementasikan menggunakan perangkat lunak Orange3. Hasil dari model Klasifikasi Naïve Bayes untuk prediksi status curah hujan di kabupaten Kebumen dan memberikan hasil yang diukur dalam nilai presisi, recall, dan akurasi, berturut-turut nilainya 0.638 , 0.653, dan 0.653. Nilai akurasi yang berada di Tengah-tengah antara decision tree yang bernilai 0.583 dan *Neural Network* bernilai 0,704 kinerja Naïve Bayes lebih baik daripada decision tree dan lebih buruk dari *Neural Network*. Hasil dari penelitian ini bermanfaat untuk diterapkan pada situasi nyata, untuk dapat memprediksi perkiraan hujan di Kabupaten Kebumen agar masyarakat lebih dapat berjaga-jaga saat musim hujan. Apabila jumlah dataset nyata dan berukuran besar dengan proporsi nilai setiap variabel yang seimbang sangat baik untuk mendapatkan akurasi prediksi yang lebih tinggi.

Referensi

- [1] S. Setiyaris, M. A. Hariyadi, and C. Crysdian, "Prediksi Curah Hujan Bulanan Berdasrkan Parameter Cuaca Menggunakan Jaringan Saraf Tiruan Levenberg Marquardt," *J. Media Inform. Budidarma*, vol. 7, no. 3, p. 1125, 2023, doi: 10.30865/mib.v7i3.6328.
- [2] R. and others Sukri, Ahmad Syarif and Sriwati, Meny and Hafid, Hendra and Serang, *Hidrologi. TOHAR MEDIA*, 2024.
- [3] S. Indratmoko, D. Harmantyo, and E. Kusratmoko, "Variabilitas curah hujan di Kabupaten Kebumen," *J. Geogr. Lingkung. Trop.*, vol. 1, no. 1, 2017, doi: 10.7454/jglitrop.v1i1.5.
- [4] A. U. Azmi, A. F. Hadi, D. Anggraeni, and A. Riski, "Naïve bayes methods for rainfall prediction classification in Banyuwangi," *J. Phys. Conf. Ser.*, vol. 1872, no. 1, 2021, doi: 10.1088/1742-6596/1872/1/012028.
- [5] S. S. and others Haris, S Pd and Nurfaika, S Si and Nurul Ainun Tangge, S and Najma Nur Mawaddah, S Pd and Boi Herman, S Si and Rusiyah, S Pd and Natalia Adel, HN and Mirna Yunita, M Pd and Syamsunardi, S Pd and Saban Rahim, *Pengantar Geografi*. PT Penerbit Qriset Indonesia, 2025.

- [6] V. I. Yani, A. Aradea, and H. Mubarak, “Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naïve Bayes dan C4.5,” *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 3, pp. 607–619, 2022, doi: 10.28932/jutisi.v8i3.5455.
- [7] I. Daqiqil, *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*. Unri Press, 2021.
- [8] M. and others Muhammad Wali, ST and Efitra, S and Kom, M and Sudipa, I Gede Iwan and Kom, S and Heryani, Ani and Sos, S and Hendriyani, Chandra and Rakhmadi Rahman, ST and Kom, *Penerapan & Implementasi Big Data di Berbagai Sektor (Pembangunan Berkelanjutan Era Industri 4.0 dan Society 5.0)*. PT. Sonpedia Publishing Indonesia, 2023.
- [9] D. Y. Liliana, H. Maulana, and A. Setiawan, “Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes,” *Multinetics*, vol. 7, no. 1, pp. 48–53, 2021, doi: 10.32722/multinetics.v7i1.3786.
- [10] E. D. Apridiansyah, Yovi and Veronika, Nuri David Maria and Putra, “Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode Naive Bayes,” *J. Sci. Appl. Informatics*, vol. 4, pp. 236--247, 2021.
- [11] D. Aryanti, “Analisis Sentimen Ibukota Negara Baru Menggunakan Metode Naïve Bayes Classifier,” *J. Inf. Syst. Res.*, vol. 3, no. 4, pp. 524–531, 2022, doi: 10.47065/josh.v3i4.1944.