

Peramalan Curah Hujan di Kabupaten Tuban Menggunakan Algoritma KNN

Vita Fatimah¹, Ahmad Zaenal Arifin², Suzatmo Putro³

^{1,2,3}Universitas PGRI Ronggolawe

³Badan Pusat Statistik

¹vitafatimah572@gmail.com

ABSTRAK

Penelitian ini bertujuan memprediksi curah hujan menggunakan data historis dan terkini dengan menerapkan metode K-Nearest Neighbors (KNN). Prediksi tersebut bermanfaat untuk mendukung berbagai aktivitas sosial-ekonomi, termasuk keselamatan publik, sektor pertanian, perkebunan, perikanan, dan penerbangan. Metode penelitian mengacu pada studi literatur dengan memanfaatkan data publikasi Badan Pusat Statistik (BPS) Kabupaten Tuban. Hasil analisis menunjukkan nilai RMSE 36,3942 dan R-square 0,6214 berdasarkan 36 sampel data bulanan yang mencakup variabel suhu, tekanan udara, kecepatan angin, kelembapan udara, dan curah hujan.

Kata Kunci: Prediksi, Curah hujan, KNN, RMSE, R-Square.

ABSTRACT

This study aims to predict rainfall using historical and current data by applying the K-Nearest Neighbors (KNN) method. These predictions are useful for supporting various socio-economic activities, including public safety, agriculture, plantations, fisheries, and aviation. The research method refers to literature studies using data published by the Central Statistics Agency (BPS) of Tuban Regency. The analysis results show an RMSE value of 36.3942 and an R-square of 0.6214 based on 36 monthly data samples covering variables such as temperature, air pressure, wind speed, air humidity, and rainfall.

Keywords: Prediction, Rainfall, KNN, RMSE, R-Square.

PENDAHULUAN

Indonesia merupakan negara tropis, yang saat ini perubahan iklimnya tidak menentu. Ketidakpastian pada perubahan iklim tersebut mengakibatkan perkiraan cuaca semakin sulit. kondisi cuaca dapat dipengaruhi oleh berbagai faktor, diantaranya seperti suhu, tekanan udara, kecepatan angin, kelembapan udara serta curah hujan (Yusuf dkk., 2022). Sementara itu, Curah hujan dapat dipengaruhi oleh suhu, tekanan udara, kecepatan angin, serta kelembapan udara. Dengan suhu yang tinggi dapat meningkatkan evaporasi, yang kemudian evaporasi tersebut dapat meningkatkan kelembapan udara. Jika kelembapan udara tinggi dan kondisi atmosfer mendukung, maka curah hujan berpotensi meningkat. Tekanan udara yang rendah sering menyebabkan pembentukan awan hujan, sedangkan kecepatan angin menentukan distribusi uap air di atmosfer, yang pada akhirnya mempengaruhi besarnya curah hujan serta lokasi curah hujan yang turun ke permukaan tanah.

Curah hujan megacu pada banyaknya air hujan yang jatuh pada permukaan tanah, yang dihitung berdasarkan volume air per satuan luas. Besarnya curah hujan yang terjadi tidak dapat ditentukan secara pasti, namun dapat diperkirakan atau diramalkan. Peramalan merupakan suatu metode prediksi nilai di masa depan dengan menganalisis data historis maupun kondisi terkini (Riadi 2020). Prediksi curah hujan dapat mendukung berbagai aktivitas sosial-ekonomi di Indonesia, dimana hasilnya bermanfaat sebagai informasi yang penting bagi beragam sektor kehidupan seperti keselamatan publik, sektor pertanian, perkebunan, perikanan, penerbangan, dan berbagai bidang lainnya (Nita puspitasi, 2023).

Metode peramalan yang dapat diterapkan dalam curah hujan ada beberapa metode, diantaranya Regresi Linear Berganda (Wicaksono dkk., 2023), Jaringan Syaraf Tiruan (JST) (Nailah dkk., 2024), Fuzzy Time Series (Safitri dkk., 2023), dan *K-Nearest Neighbors* (KNN) (Sofiyah and Azkiyah 2022). Beberapa metode diatas memiliki kelemahan yang berbeda dalam mengerjakan suatu masalah. Dalam penggunaan metode Regresi Linear Berganda tidak dapat menangkap hubungan non-linear antar variabel input dengan variabel output, pada metode JST sering memerlukan waktu yang lebih lama dalam penggerjaannya dibanding metode yang lain, untuk metode Fuzzy Time Series dalam penggerjaannya kurang efektif untuk pola data yang acak, dan metode KNN sensitif terhadap skala data, maka diperlukan adanya standarisasi.

Kelebihan dari algoritma KNN dalam meramalkan curah hujan adalah KNN dapat digunakan untuk berbagai jenis data, yaitu data numerik maupun data kategorikal (Rahayu dkk., 2024). KNN adalah suatu algoritma *machine learning* yang tergolong ke dalam *supervised learning*. Pada dasarnya, algoritma KNN diaplikasikan untuk mengklasifikasi objek dengan memanfaatkan data *training* yang memiliki selisih nilai minimal dengan memperhitungkan jarak terdekat antar tetangga atau anggota (Maskuri dkk., 2022).

Adapun beberapa penelitian terdahulu yang membahas metode KNN tentang curah hujan, antara lainnya adalah (Rofiq dkk., 2020) didapatkan hasil prediksi dengan $k=3$ mendapatkan nilai RMSE sebesar $9.899+/-0.000$. (Nanda dkk., 2022) didapatkan hasil prediksi dengan $k=5$ dengan pengujian akurasi menggunakan Confusion Matrix diperoleh 84.38%. (Agung dkk., 2023) dengan $k=3$ didapatkan akurasi mencapai 82,21%. Dengan adanya beberapa penelitian tersebut peneliti menggunakan metode KNN dikarenakan metode ini mampu menangani data dengan struktur penggerjaan yang mirip dengan *clustering*, tidak seperti metode lainnya, serta metode KNN ini mudah dalam menangani data non-linear seperti curah hujan.

Pada pembahasan ini menggunakan algoritma KNN dalam meramalkan besarnya curah hujan yang akan datang. Algoritma KNN ini mampu memprediksi data curah hujan dengan baik menggunakan metode berbasis jarak yang diterapkan dalam data *training* dan data *testing*. Metode berbasis jarak yang diterapkan adalah jarak *Euclid*, untuk mengolah data dan melihat hasil dari prediksi tersebut. Untuk mengetahui hasil kinerja dari algoritma KNN dapat menggunakan *root mean square error* (RMSE) untuk mengetahui *error* dan *R Square* untuk mengetahui hubungan antara variabel suhu, tekanan udara, kecepatan angin, serta kelembapan udara dan varibel curah hujan.

METODE PENELITIAN

Pada penelitian ini dilakukan pengumpulan data dengan Studi Kepustakaan (*library research*) di website Badan Pusat Statistik, dan mempelajari beberapa artikel ilmiah, serta hasil penelitian sebelumnya. Pengumpulan data berupa data sekunder yang telah ada pada publikasi di website BPS Kabupaten Tuban. Data yang ada dalam penelitian ini adalah data sekunder, dimana data diperoleh dari publikasi Kabupaten Tuban Dalam Angka pada tahun 2022 – 2024 dari Badan Pusat Statistik Kabupaten Tuban dengan jumlah 36 data dengan variabel suhu, tekanan udara, kecepatan angin, serta kelembapan udara, dan curah hujan. Dalam meramalkan curah hujan menggunakan metode KNN.

Standarisasi Data

Standarisasi data adalah suatu proses mengganti data mentah menjadi satuan pengukuran yang seragam pada data yang digunakan. standarisasi data dapat memudahkan dalam proses melakukan analisis dalam visualisasi data (Sudipa dkk., 2023). Standarisasi data dilakukan ketika terdapat perbedaan skala antar variabel input yang dapat menyebabkan bias dalam hasil. Oleh karena itu, penerapan standarisasi terhadap data menjadi langkah penting yang harus dilakukan sebelum adanya proses perhitungan pada metode yang diterapkan (Nisrina dkk., 2022).

Standarisasi diterapkan dengan mengubah data menjadi nilai *z-score* untuk membandingkan data dari distribusi yang berbeda (Iskandar and Latifa 2023). Berikut ini merupakan rumus yang digunakan dalam standarisasi data

$$z = \frac{(x - \mu)}{\sigma}$$

dengan:

z	Nilai data yang telah di standarisasi (<i>z-score</i>)
x	Nilai data asli
μ	Rata-rata (<i>mean</i>) dari seluruh dataset
σ	Standar deviasi dari seluruh dataset

dimana interpretasi *z-score* adalah

- $z = 0$ Data berada tepat pada rata-rata
- $z > 0$ Data berada di atas rata-rata.
- $z < 0$ Data berada di bawah rata-rata.

Algoritma K-Nearest Neighbour (KNN)

Algoritma K-Nearest Neighbor (KNN) merupakan metode klasifikasi yang memprediksi kategori objek baru berdasarkan sejumlah (K) tetangga terdekatnya. Prinsip kerja algoritma ini adalah mengelompokkan objek baru dengan membandingkan karakteristiknya terhadap sampel data latih. Sebagai bagian dari Supervised Learning, KNN menentukan kelas instance baru melalui voting mayoritas label kelas dari tetangga terdekatnya, dimana kelas yang paling dominan akan ditetapkan sebagai hasil klasifikasi (Putri 2021).

Dalam implementasinya, perhitungan jarak antar objek dapat dilakukan menggunakan Euclidean Distance. Metode ini mengkuantifikasi tingkat kedekatan antara dua objek dalam ruang multidimensi. Rumus Euclidean Distance dinyatakan sebagai:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (x_{ir} - x_{jr})^2}$$

dengan:

$d(x_i, x_j)$	Jarak data <i>testing</i> ke- i dengan data <i>training</i> ke- j
x_{ir}	Data <i>training</i> ke- i pada variabel ke- r
x_{jr}	Data <i>testing</i> ke- j pada variabel ke- r
n	Dimensi data variabel bebas
r	Variabel data

Mekanisme kerja KNN melibatkan tiga komponen utama: data latih, data uji, dan nilai k. Prosesnya diawali dengan pengurutan data berdasarkan jarak terdekat antara data uji dan latih, kemudian menentukan output berdasarkan rata-rata nilai k tetangga terdekat. Secara rinci, tahapan perhitungan KNN adalah sebagai berikut (Rahmat Dian Nugraha dkk., 2020).

- Penentuan parameter k
- Perhitungan jarak antara data evaluasi dengan semua data *training*
- Pengurutan hasil perhitungan jarak
- Seleksi k jarak terdekat
- Pemetaan kelas sesuai dengan tetangga terdekat
- Penentuan kelas akhir berdasarkan mayoritas kelas tetangga

HASIL DAN PEMBAHASAN

Berikut adalah langkah yang terapkan dalam penggerjaan metode KNN:

A. Pembagian data

Data yang diambil adalah data curah hujan bulanan di Kabupaten Tuban pada 3 tahun terakhir selama periode 2022 – 2024 dengan 4 variabel X dan 1 variabel Y dengan jumlah 36 data. Data tersebut diperoleh dari publikasi BPS Kabupaten Tuban. Setelah pengumpulan data dilakukan, langkah selanjutnya yaitu pembagian data, pada pembagian data ini dibantu oleh pemograman *python* pada *website google colab*. Pembagian data dibagi menjadi 2 yaitu data *training* dan data *testing* dengan proporsi 70:30, data *training* terdapat 25 data dan data *testing* terdapat 11 data.

Tabel 1 data *training*

No.	Data ke-	Suhu (X ₁)	Kelembaban (X ₂)	Kecepatan Angin (X ₃)	Tekanan Udara (X ₄)	Curah Hujan (Y)
1	1	26.5	87	12	1006.8	323.2
2	2	26.7	86	16	1005.8	93.3
3	3	27.3	84	10	1007	218.3
4	4	27.7	81	11	1007.75	77.1
5	5	28.3	80	10	1008.05	49.4

6	6	27.8	76	10	1008.9	166
7	7	27.3	66	8	1009.25	4.3
8	8	27.9	55	10	1009.6	13.7
9	11	27.9	71	10	1005.95	333.1
10	12	27.6	72	11	1006.85	278.7
11	15	27.5	86	12	1006	165.3
12	16	28	84	8.7	1006.9	131.5
13	19	27.3	81	13	1007.55	126.7
14	20	27.5	79	17	1008.1	46.2
15	21	28	77	14.7	1007.75	47.8
16	23	27.5	84	10.2	1006.9	171.1
17	24	27.2	83	13	1006.2	296
18	25	27.1	84	8.64	1006.6	227.1
19	26	26.7	85	10.2	1006.95	113.8
20	28	28	83	5.48	1007.4	134.6
21	29	27.9	79	6.4	1007.8	76.5
22	30	28.1	77	6.95	1008.65	4.6
23	33	27.8	69	9.42	1009.45	0.4
24	34	29.2	71	9.35	1009.1	0
25	35	29.7	73	8.08	1007.3	60.5

Tabel 2 data *testing*

No.	Data ke-	X ₁	X ₂	X ₃	X ₄	Y
1	9	28.3	68	12	1008.95	81.5
2	10	28.7	63	11	1008.4	92.5
3	13	27.1	85	15	1006.9	111.3
4	14	26.9	86	14.5	1006.6	247.1
5	17	28.2	85	11	1007.25	138
6	18	27.4	83	11	1006.9	112.7
7	22	27.6	80	13	1007.5	171.6
8	27	27.4	84	5.72	1008.5	133.1
9	31	27.4	73	7.64	1009.45	16.9
10	32	27.2	74	9.48	1010.8	40.9
11	36	28.9	76	6.77	1008.25	99.4

B. Standarisasi data

Pada bagian standarisasi ini dilakukan pada variabel X dalam data *training* dan data *testing* untuk memastikan data telah memiliki skala yang sama, sehingga pada penggerjaan metode dapat dilakukan dengan lebih efisien. Standarisasi data dibantu dalam penggerjaannya menggunakan pemograman *python* pada *website google colab* menggunakan persamaan (1). Setelah standarisasi dilakukan dapat dibentuk model algoritma KNN dalam melakukan prediksi.

Tabel 3 standarisasi data *training*

No.	Data ke-	X ₁	X ₂	X ₃	X ₄
1	1	-1.74371458e+00	1.18891803e+00	5.71854956e-01	-6.76625390e-01
2	2	-1.45309548e+00	1.05503086e+00	2.04267532e+00	-1.58606812e+00
3	3	-5.81238194e-01	7.87256530e-01	-1.63555224e-01	-4.94736844e-01
4	4	5.16243220e-15	3.85595035e-01	2.04149866e-01	1.87345202e-01
5	5	8.71857291e-01	2.51707870e-01	-1.63555224e-01	4.60178021e-01
6	6	1.45309548e-01	-2.83840790e-01	-1.63555224e-01	1.23320434e+00
7	7	-5.81238194e-01	-1.62271244e+00	-8.98965404e-01	1.55150929e+00
8	8	2.90619097e-01	-3.09547126e+00	-1.63555224e-01	1.86981425e+00
9	11	2.90619097e-01	-9.53276615e-01	-1.63555224e-01	-1.44965171e+00
10	12	-1.45309548e-01	8.19389450e-01	-2.04149866e-01	-6.31153254e-01
11	15	-2.90619097e-01	1.05503086e+00	5.71854956e-01	-1.40417957e+00
12	16	4.35928645e-01	7.87256530e-01	-6.41571841e-01	-0.573847956
13	19	-5.81238194e-01	3.85595035e-01	9.39560046e-01	5.45665637e-03
14	20	-2.90619097e-01	1.17820705e-01	2.41038041e+00	5.05650157e-01
15	21	4.35928645e-01	-1.49953625e-01	1.56465870e+00	1.87345202e-01
16	23	-2.90619097e-01	7.87256530e-01	-9.00142060e-02	-5.85681117e-01
17	24	-7.26547742e-01	6.53369365e-01	9.39560046e-01	-1.22229103e+00
18	25	-8.71857291e-01	7.87256530e-01	-6.63634146e-01	-8.58513936e-01
19	26	-1.45309548e+00	9.21143695e-01	-9.00142060e-02	-5.40208981e-01
20	28	4.35928645e-01	6.53369365e-01	-1.82558223e+00	-1.30959753e-01
21	29	2.90619097e-01	1.17820705e-01	-1.48729355e+00	2.32817339e-01
22	30	5.81238194e-01	-1.49953625e-01	-1.28505575e+00	1.00584366e+00
23	33	1.45309548e-01	-1.22105095e+00	-3.76824176e-01	1.73339784e+00
24	34	2.17964323e+00	-9.53276615e-01	-4.02563532e-01	1.41509289e+00
25	35	2.90619097e+00	-6.85502285e-01	-8.69548997e-01	-2.21904026e-01

Pada standarisasi data *training* merepresentasikan bahwa nilai dalam notasi ilmiah. Representasi notasi ilmiah sama halnya dengan representasi desimal, hanya saja berbeda dalam segi penulisan.

Tabel 3 standarisasi data *testing*

No.	Data ke-	X ₁	X ₂	X ₃	X ₄
1	9	0.87185729	-1.35493811	0.57185496	1.27867648
2	10	1.45309548	-2.02437394	0.20414987	0.77848298
3	13	-0.87185729	0.9211437	1.67497023	-0.58568112
4	14	-1.16247639	1.05503086	1.49111768	-0.85851394
5	17	0.72654774	0.9211437	0.20414987	-0.26737616
6	18	-0.43592865	0.65336937	0.20414987	-0.58568112
7	22	-0.14530955	0.25170787	0.93956005	-0.04001548

8	27	-0.43592865	0.78725653	-1.73733301	0.86942725
9	31	-0.43592865	-0.68550228	-1.03133924	1.73339784
10	32	-0.72654774	-0.55161512	-0.35476187	2.96114552
11	36	1.74371458	-0.28384079	-1.35124266	0.64206657

Pada standarisasi data *testing* merepresentasikan desimal, pada representasi desimal dan representasi notasi ilmiah itu pada dasarnya sama, akan tetapi ditulis dalam format yang berbeda.

C. Algoritma KNN

Berdasarkan standarisasi data *training* dan data *testing* pada variabel X, langkah selanjutnya yaitu pembentukan model KNN menggunakan jarak *euclid* pada persamaan (2) dengan nilai k=8, dalam pemilihan nilai k dibantu dengan pemograman *python* pada website *google colab*.

Dibawah ini merupakan perhitungan data *testing* pertama yang dihitung dengan data *training* yang ada menggunakan jarak *Euclid*.

$$d_1 = \sqrt{(-1.74371458 - 0.87185729)^2 + (1.18891803 - (-1.35493811))^2 + (0.571854956 - 0.57185496)^2 + (-0.67662539 - 1.27867648)^2} \\ = 4.139519981$$

Setelah didapatkan semua hasil jarak *Euclid* dari data *testing* pertama hingga data *testing* terakhir, langkah selanjutnya yaitu mengurutkan hasil jarak dari yang terkecil hingga terbesar, dan didapatkan peringkat pada setiap jarak berdasarkan jarak pada data *testing*.

Tabel 1 Hasil Jarak *Euclid* pada data *testing* 1 berdasarkan peringkat

Data ke-	Hasil Jarak Euclid Data <i>Testing 1</i>	Peringkat
33	1.285519744	1
6	1.489300824	2
34	1.685151879	3
5	1.947326159	4
21	1.954145495	5

Setelah mengurutkan peringkat dari data *testing* 1 hingga data *testing* terakhir, dapat dilanjurkan dengan mengurutkan hasil jarak sesuai dengan nilai k pada semua data *testing* dan juga data *training*, pada hal ini diambil nilai k=8, dapat dilihat dibawah ini hasil pengurutan nilai k=8 pada data testing 1.

Tabel 2 Hasil Jarak *Euclid* pada data *testing* 1 berdasarkan peringkat dengan k=8

Data ke	Hasil Jarak Euclid Data <i>Testing No.1</i>	Peringkat
33	1.285519744	1
6	1.489300824	2
34	1.685151879	3
5	1.947326159	4

21	1.954145495	5
8	2.063386897	6
7	2.102603116	7
30	2.249222577	8

Setelah mengurutkan setiap data testing yang ada dengan $k=8$, dari langkah ini dapat dilanjut dengan mencari hasil dari prediksi yang akan dicari.

D. Hasil Prediksi

Setelah diurutkan berdasarkan peringkat tertinggi dari setiap data *testing* dari data *testing* pertama hingga data *testing* terakhir, maka selanjutnya dapat dilakukan prediksi. Untuk menentukan hasil prediksi dengan cara mencari rata-rata curah hujan dari tiap data tersebut. Maka didapatkan hasil seluruh prediksi yang dapat dilihat pada tabel 3.

Tabel 3 Nilai Aktual dan Hasil Prediksi

No.	Data ke-	Nilai Aktual	Hasil Prediksi
1	9	81.5	35.775
2	10	92.5	71.625
3	13	111.3	151.5375
4	14	247.1	188.4625
5	17	138	123.4
6	18	112.7	174.425
7	22	171.6	157.7375
8	27	133.1	129.4625
9	31	16.9	64.1125
10	32	40.9	49
11	36	99.4	77.8875

E. Pengujian Model

Didapatkan hasil RMSE sebesar 36.3942 menunjukkan bahwa tingkat RMSE yang dihasilkan dari $k=8$ adalah sebesar 36.3942. Berbeda dengan penelitian sebelumnya yang dilakukan oleh (Rofiq dkk., 2020) mendapatkan hasil RMSE sebesar $9.899+/-0.000$ perbedaan hasil dengan yang peniliti lakukan disebabkan oleh dataset yang dipakai dalam meramalkan curah hujan terbatas, yang hanya menggunakan data bulanan pada rentang waktu 3 tahun dengan 36 data, berbeda dengan penelitian sebelumnya yang menggunakan data harian dengan jumlah 335 data, dikarenakan metode KNN ini merupakan algoritma yang bergantung pada jarak antar titik data dengan parameter k , maka semakin banyak data, akan semakin baik pula pengolahan dalam metode KNN, akan tetapi pada penelitian sebelumnya tidak menjelaskan hubungan antar variabel, dan pada penelitian ini menjelaskannya dan didapatkan nilai *R-Square* sebesar 0.6214 yang berarti curah hujan di Kabupaten Tuban yang dipengaruhi oleh 4 variabel yaitu, suhu, kelembaban, kecepatan angin, dan tekanan udara sebesar 0.6214 secara rata-rata menunjukkan bahwa model KNN cukup baik dalam menjelaskan hubungan antara variabel X dalam meramalkan curah hujan

dan sisanya sebesar 0,3786 dipengaruhi oleh variabel lain yang tidak dimasukkan dalam model.

PENUTUP

Berdasarkan hasil analisis algoritma KNN yang telah diterapkan dalam meramalkan curah hujan dapat disimpulkan bahwa hasil prediksi memperoleh *Root Mean Square Error* (RMSE) sebesar 36.3942, hasil RMSE ini disebabkan oleh dataset yang dipakai dalam meramalkan curah hujan terbatas, dan hasil R-Square sebesar 0.6214 yang berarti curah hujan di Kabupaten Tuban yang dipengaruhi oleh 4 variabel didapatkan hasil sebesar 0.6214 secara rata-rata, menunjukkan bahwa model KNN cukup baik dalam menjelaskan hubungan antara variabel X dalam meramalkan curah hujan dan sisanya sebesar 0,3786 dipengaruhi oleh variabel lain yang tidak dimasukkan dalam model.

REFERENSI

- Agung, Andi Sadri, A. Arfan Fauzi, Andi Akram Nur Risal, and Fhatiah Adiba. 2023. “Implementasi Teknik Data Mining Terhadap Klasifikasi Data Prediksi Curah Hujan BMKG Di Sulawesi Selatan.” *Jurnal Tekno Insentif* 17(1):22–23. doi: 10.36787/jti.v17i1.955.
- Iskandar, Mohammad Amirulhaq, and Ulinnuha Latifa. 2023. “Website Prediksi Customer Churn Untuk Mempertahankan Pelanggan Pada Perusahaan Telekomunikasi.” *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(2):1308–16.
- Maskuri, Muhammad Naja, Kadek Sukerti, and R. M. Herdian Bhakti. 2022. “Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Penyakit Stroke Stroke Disease Predict Using KNN Algorithm.” *Jurnal Ilmiah Intech : Information Technology Journal of UMUS* 4(1):130–40.
- Nailah, Fadia, Dwi Ina Larasati, Siswanto Siswanto, and Anisa Kalondeng. 2024. “Optimasi Metode Jaringan Saraf Tiruan Backpropagation Untuk Peramalan Curah Hujan Bulanan Di Kota Denpasar.” *MATHunesa: Jurnal Ilmiah Matematika* 12(1):134–40. doi: 10.26740/mathunesa.v12n1.p134-140.
- Nanda, Deden Martia, Tacbir Hendro Pudjiantoro, and Puspita Nurul Sabrina. 2022. “Metode K-Nearest Neighbor (KNN) Dalam Memprediksi Curah Hujan Di Kota Bandung.” *Seminar Nasional Teknik Elektro, Sistem Informasi, Dan Teknik Informatika* 387–93.
- Nisrina, Sausan, Wiwit Pura Nurmayanti, and Muhammad Gazali. 2022. “Penerapan Metode Clustering SOM Dan DBSCAN Dalam Mengelompokkan Unmet Need Keluarga Berencana Di Nusa Tenggara Barat.” *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika* 15(2):237–44.
- NITA PUSPITASARI, NITA PUSPITASARI. 2023. “Analisis Debit Banjir Menggunakan Metode Mean Annual Flood Di DAS Majene Kabupaten Majene.”
- Putri, Ayu Azlina. 2021. “Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus : PT . Central Brastagi Utama).” 1(6):354–61.
- Rahayu, Prastyadi Wibawa, I. Gede Iwan Sudipa, Suryani Suryani, Arie Surachman, Achmad Ridwan, I. Gede Mahendra Darmawiguna, Muh Nurtanzis Sutoyo, Isnandar Slamet, Sitti Harlina, and I. Made Dendi Maysanjaya. 2024. *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia.
- Rahmat Dian Nugraha, Adhitya, Karina Auliasari, and Yosep Agus Pranoto. 2020. “IMPLEMENTASI METODE K-NEAREST NEIGHBOR (KNN) UNTUK

- SELEKSI CALON KARYAWAN BARU (Studi Kasus : BFI Finance Surabaya)." *JATI (Jurnal Mahasiswa Teknik Informatika)* 4(2):14–20. doi: 10.36040/jati.v4i2.2656.
- Riadi, Hafid. 2020. "Proyeksi Permintaan Minyak Bumi Di Tengah Kasus Covid-19."
- Rofiq, Harun, Kartika Chandra Pelangi, and Yuliyanti Lasena. 2020. "Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma Naive Bayes." *Jurnal Manajemen Informatika Dan Sistem Informasi* 3(1):8–15.
- Safitri, Kurniawati, Dadan Kusnandar, and Naomi Nessyana Debataraja. 2023. "Peramalan Curah Hujan Dengan Metode Fuzzy Time Series Markov Chain." *Buletin Ilmiah Math, Stat, Dan Terapannya (BIMASTER)* 12(1):35–42.
- Sofiyan, Amat, and Ahmedika Azkiya. 2022. "Penerapan Metode Rough Set Menganalisis Penyakit Yang Sering Dikeluhkan Pasien (Studi Kasus Puskesmas Jaya Mukti Dumai)." *Informatika* 14(1):31–40.
- Sudipa, I. Gede Iwan, Ida Bagus Gede Sarasvananda, Heri Prayitno, I. Nyoman Tri Anindia Putra, Risanto Darmawan, and Dwi Atmodjo WP. 2023. *Teknik Visualisasi Data*. PT. Sonpedia Publishing Indonesia.
- Wicaksono, Riyanto, Joseph Dedy Irawan, and Suryo Adi Wibowo. 2023. "Sistem Peramalan Curah Hujan Menggunakan Metode Regresi Linier Berganda Berbasis IoT." 7(4):2103–12.
- Yusuf, Muhammad, Arief Setyanto, and Komang Aryasa. 2022. "Analisis Prediksi Curah Hujan Bulanan Wilayah Kota Sorong Menggunakan Metode Multiple Regression." *Jurnal Sains Komputer & Informatika (J-SAKTI)* 6(1):405–17.