Epsilon: Jurnal Matematika Murni dan Terapan
http://ppjp.ulm.ac.id/journals/index.php/epsilon
p-ISSN: 1978-4422 | e-ISSN 2656-7660
Vol.17 No.2 (Desember 2023) Hal. 186-199

IMPLEMENTASI *K-NEAREST NEIGHBOR* DALAM MEMPREDIKSI CURAH HUJAN GUNA MENYUSUN PENJADWALAN TANAM PADI DI PULANG PISAU

Yuniarta Basani[1], Dian Nur Handayani[2], Herman Santoso Pakpahan[3], Regina Wahyudyah Sonata Ayu[1]

[1]Program Studi Matematika, Fakultas MIPA, Universitas Palangka Raya
[2]Program Studi Informatika, Fakultas Teknik, Universitas Mulawarman
[3]Program Studi Sistem Informasi, STMIK Palangka Raya
Email: *yuniartabasani@mipa.upr.ac.id*

ABSTRACT

Rice is a crucial commodity in both social and economic life in Indonesia, as it serves as the main source of carbohydrates. The Climate change affects planting patterns in the agricultural sector and rainfall plays a significant role in plant growth as it affects the availability of water. The district of Pulang Pisau in Central Kalimantan has excellent agricultural potential, especially in rice production. This study aims to predict rainfall using the K-NN method and convert the predicted rainfall results into a crop calendar table to minimize crop failure. In recent decades, the K-Nearest Neighbor (K-NN) method has been used not only for classification but also for prediction. The K-NN method was implemented in this study for the prediction of rainfall, with its performance tested using RMSE. Data from the official BMKG website for the years 2011-2022 was used. The data was divided into training and testing sets in ratios of 70:30, 80:20, and 90:10, respectively. The testing results for the K-NN method in predicting rainfall yielded the smallest RMSE value of 109.3, calculated using Manhattan Distance with a value of k = 16. The rain forecast data, converted to a crop calendar structure, indicates that rice cultivation can be carried out between July, August, and November 2023, also June 2024.

Keywords: data mining, K-NN, prediction, schedule for planting

Received: 27 Oktober 2023, Accepted: 6 November 2023, Published: 1 Desember 2023

ABSTRAK

Beras merupakan komoditas penting dalam kehidupan sosial dan ekonomi di Indonesia serta menjadi sumber karbohidrat utama. Perubahan iklim mempengaruhi pola tanam dalam sektor pertanian, dan curah hujan memiliki peran penting dalam pertumbuhan tanaman karena mempengaruhi ketersediaan air. Kabupaten Pulang Pisau di Kalimantan Tengah memiliki potensi pertanian yang baik, terutama dalam produksi padi. Untuk meminimalisir kegagalan panen akibat perubahan iklim, prediksi curah hujan diperlukan sebagai acuan dalam penjadwalan tanam padi. Penelitian ini bertujuan untuk memprediksi curah hujan menggunakan metode K-NN dan mengkonversikan hasil prediksi curah hujan menjadi tabel kalender tanam guna meminimalisir terjadinya kegagalan panen. Metode K-NN pada penelitian ini diimplementasikan dalam memprediksi curah hujan, dengan pengujian performa metode K-NN menggunakan metode RMSE. Data yang digunakan berasal dari website resmi BMKG yang diambil dari Tahun 2011-2022. Rasio pembagian *data training* dan *data testing* yang digunakan adalah 70:30, 80:20 dan 90:10. Hasil pengujian penerapan metode K-NN dalam prediksi curah hujan menghasilkan nilai RMSE terkecil 109,3 dengan menggunakan perhitungan jarak *Manhattan Distance* dan nilai k = 16. Hasil prediksi curah hujan yang dikonversikan ke dalam kalender tanam menunjukkan bahwa Juli, Agustus, dan November 2023, serta Juni 2024 dapat dilakukan proses penanaman padi.

Kata kunci: *data mining*, jadwal tanam, K-NN, prediksi

PENDAHULUAN

Beras adalah salah satu komoditas penting dalam kehidupan sosial ekonomi sekaligus sumber daya pangan karbohidrat utama dalam stuktur pangan masyarakat Indonesia. Proses produksi pertanian membutuhkan keadaan iklim yang baik. Kondisi iklim yang tidak menentu dapat menurunkan hasil produksi pertanian serta dapat membuat petani mengalami gagal panen. Salah satu unsur iklim yang berpengaruh dalam menentukan pola tanam adalah curah hujan. Curah hujan berperan bagi pertumbuhan dan perkembangan tanaman karena berpengaruh terhadap ketersediaan air (Cahyati, 2019).

Petani pada umumnya menanam padi sekitar bulan November sampai Maret. Namun pada tahun 2021, petani pulang pisau mengalami gagal tanam dan produksi padi terancam menurun (Zona kalteng, 2020). Bentuk upaya memaksimalkan produksi pertanian tanaman padi adalah dengan memprediksi curah hujan. Prediksi curah hujan dilakukan di Kabupaten Pulang Pisau yang menjadi daerah produksi padi terbesar kedua di Provinsi Kalimantan Tengah. Kabupaten Pulang Pisau juga menjadi tujuan transmigrasi pengembangan *Food Estate* oleh Pemerintah Pusat (Evi Fitriana, 2021).

Beberapa peneliti telah melakukan prediksi menggunakan *data mining*. Pada tahun 2022, Faidah memprediksi curah hujan menggunakan metode *North American Multi Model Ensemble* dengan memperoleh nilai RMSE (*Root Mean Square Error*) sebesar 148,6 (Faidah, 2022). Pada tahun 2023, Laily memprediksi curah hujan menggunakan metode *Fuzzy Time Series-Markov Chain* dan menghasilkan MAPE (*Mean Absolute Percentage Error*) kurang dari 10% (Laily et al., 2023). Pada tahun yang sama, Naisya memprediksi suhu awan dan hujan menggunakan K-NN (*K-Nearest Neighbor*) dan memperoleh akurasi pengujian terendah 40-50% dan tertinggi 90-100% (Nisya, 2023). Metode K-NN memiliki kelebihan yaitu *Asymptotically correct* dan dapat mengeneralisasi himpunan data training yang relatif kecil (Rasyid, 2022). Berdasarkan penelitian tersebut, penulis melakukan penelitian menggunakan K-NN untuk memprediksi curah hujan dengan parameter yang berbeda dan data yang pendukung berupa data *time series* harian yang diperoleh dari BMKG Kalimantan Tengah.

Prediksi curah hujan menggunakan metode K-NN diharapkan dapat menjadi solusi bagi masyarakat di Kabupaten Pulang Pisau dalam menentukan penjadwalan tanam padi guna meminimalisir terjadinya gagal panen. Tujuan penelitian ini adalah untuk memprediksi curah hujan dan memberikan rekomendasi jadwal yang yang baik sesuai dengan kebutuhan curah hujan untuk melakukan tanam padi di Kabupaten Pulang Pisau.

TINJAUAN PUSTAKA

1. Data Mining

Data mining adalah serangkaian proses untuk menggali, menemukan dan menambangkan data berdasarkan informasi yang dimiliki untuk menemukan pola dan relasi (Arwansyah & Arie, 2019). Istilah lain data mining adalah *knowledge discovery in database* (KDD). KDD merupakan suatu proses yang terdiri dari serangkaian proses integrasi yang berurutan, sedangkan data mining adalah bagian dari salah satu proses yang terdapat di KDD.

Langkah-langkah dalam proses KDD adalah (Yunita & Nurahman, 2022):

a. *Data Cleaning*: menghilangkan data noise dan data tidak konsisten.
b. *Data Integration*: menggabungkan data yang berasal dari berbagai sumber data.
c. *Data Selection*: menerima data yang sesuai guna dianalisis yang akan diambil *database*.
d. *Data Transformation*: mengubah data ke dalam format yang sesuai dengan proses *data mining* dengan melakukan agregasi dan *summary*.
e. *Data Mining*: melakukan proses mining untuk mengekstrak data.
f. *Pattern Evaluation*: mengidentifikasikan pola menarik yang menggambarkan *knowledge*.
g. *Knowledge Presentation*: menampilkan pengetahuan hasil proses kepada *user*.

2. Curah Hujan

Curah hujan merupakan jumlah air hujan yang turun pada suatu daerah dan waktu tertentu. Satu millimeter curah hujan menunjukkan satu wilayah dengan luas satu meter persegi pada permukaan datar menampung air dengan tinggi satu millimeter. Tanaman padi membutuhkan curah hujan sebesar 150-200 mm/ bulan dengan lama waktu penanaman adalah 4 bulan (Cahyati, 2019). Terkhusus tanaman padi yang merupakan komoditas utama, maka tanaman padi memiliki toleransi yang diberikan jika sistem menunjukkan tanaman padi tidak dapat ditanam pada tahun ini. Toleransi tanaman padi yang diberikan hanya satu kali periode tanam. Satu kali periode tanam tersebut harus memenuhi beberapa syarat, yakni:

a. Pada bulan pertama mempunyai curah hujan lebih dari 150 mm/ bulan.
b. Tiga bulan berikutnya mempunyai curah hujan yang lebih dari 100 mm/ bulan.

3. Pre-processing Data

Pre-processing data adalah sebuah teknik yang digunakan untuk mengambil informasi dari suatu data yang masih mentah. *Pre-processing data* dilakukan untuk menghilangkan *noise* yang terdapat pada data mentah, sehingga data yang nanti

digunakan pada proses *data mining* adalah data dengan kualitas baik (Kasanah et al., 2019). Menurut Nofrizal (2021) tahapan *pre-processing data* terdiri dari:

a. *Data cleaning* digunakan untuk memperbaiki data dari format nilai yang tidak konsisten menjadi data yang dapat dianalisis.
b. *Data transformation* digunakan untuk mengubah data ke dalam bentuk numerik sehingga data tersebut dapat digunakan pada proses *data mining*. Beberapa strategi yang dapat dilakukan pada tahap ini yaitu normalization, *attribute selection*, *discretization* dan *concept hierarchy generation*.
c. Data integration merupakan proses penggabungan data dari berbagai *database* ke sebuah *database* baru. Data integration dilakukan pada atribut-atribut yang mengidentifikasi entitas unik.
d. *Data reduction* adalah proses yang dilakukan untuk mendapatkan pengurangan representasi kumpulan data yang jauh lebih di dalam volume, akan tetapi belum menghasilkan hasil yang sama yang diperoleh dari suatu data analisis.

4. *K-Nearest Neighbor*

Metode *K-Nearest Neighbor* (K-NN) memiliki prinsip kerja yang sederhana, yaitu dengan menentukan k objek yang diperoleh dari *data training* yang letaknya paling dekat datang data testing (Febriyanti, 2022). Cara kerja metode K-NN ialah mencari jarak terdekat antara data yang akan dievaluasi dengan tetangga terdekat dalam *data training* untuk mencari k tetangga terdekat, yang nantinya nilai k tetangga terdekat menjadi prediksi dari *data testing* (Nikmatun & Waspada, 2019). Langkah-langkah metode K-NN untuk menghitung nilai prediksi adalah (Rilwanu et al., 2022):

a. Menentukan nilai k dan fungsi jarak.
b. Menghitung jarak antar data yang akan diprediksi dengan data sebelumnya. Menghitung jarak antar data menggunakan rumus perhitungan jarak *Euclidean Distance* dan *Manhattan Distance*. Berikut rumus persamaan *Euclidean Distance* terdapat pada persamaan (1) dan rumus persamaan *Manhattan Distance* terdapat pada persamaan (2).

$$d(x,y) = \sqrt{\sum_{i=0}^{n}(x_i - y_i)^2} \quad (1)$$

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i| \quad (2)$$

Keterangan pada rumus jarak (1) dan (2) sebagai berikut.

d : jarak antar x dan y
x : data pusat klaster
y : data pada atribut
i : setiap data

n : jumlah data

c. Mengurutkan data-data sebelumnya berdasarkan dengan jarak terkecil, kemudian menghitung rata-rata target dari k data teratas. Hasil perhitungan rata-rata tersebut adalah hasil prediksinya. Untuk menghitung rata-rata target dari k sebagai jumlah tetangga terdekat dapat menggunakan persamaan (3).

$$y(t+1) = \frac{1}{k}\sum_{i=1}^{k} y_i \tag{3}$$

5. Root Means Squared Error

Root Means Squared Error (RMSE) menurut Wai dalam (Faidah, 2022) menyebutkan bahwa RMSE merupakan salah satu metode yang digunakan untuk evaluasi kebaikan pada suatu model peramalan. Pengukuran model peramalan bertujuan untuk mendapatkan nilai terbaik antara satu nilai dan nilai hasil peramalan seperti dikemukakan oleh Haviluddin dalam (Contena, 2022). Rumus RMSE dapat dilihat pada persamaan (4).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - yi')^2} \tag{4}$$

Keterangan perhitungan RMSE sebagai berikut.

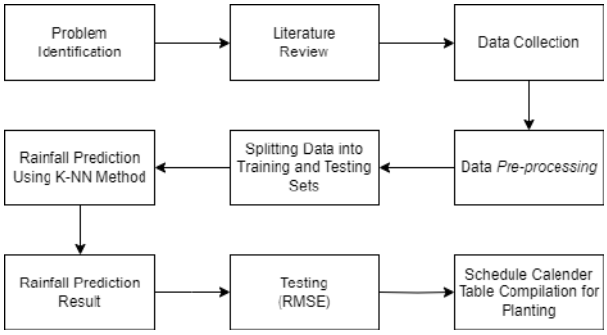
y : data aktual

y' : data hasil prediksi

n : jumlah data uji

METODE PENELITIAN

Penelitian diawali dengan pengumpulan data *time series* dengan periode bulanan terhitung dari November 2011 – Juni 2023 yang diperoleh dari *website* BMKG Provinsi Kalimantan Tengah yaitu https://dataonline.bmkg.go.id/ (BMKG, n.d.). Selanjutnya dilakukan *pre-processing* data dengan tahap *data cleaning* dan *data transformation*. Data yang telah siap diolah kemudian dibagi menjadi dua yaitu *data training* dan *data testing* untuk diimplementasikan ke dalam metode K-NN. Hasil prediksi curah hujan terhitung 1 tahun dari Juli 2023 sampai Juni 2024. Evaluasi model untuk menguji validasi data saat proses *training* dan *testing* yaitu menggunakan RMSE. Terakhir, menyusun kalender tanam padi dengan menjadikan prediksi curah hujan dan standar kebutuhan curah hujan sebagai acuannya.



Gambar 1. Metode Penelitian

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data sekunder yang diperoleh dari BMKG Provinsi Kalimantan Tengah terhitung dari November 2011 sampai Juni 2023. Sebelum dilakukan proses prediksi, data-data tersebut diolah terlebih dahulu melalui tahapan *pre-processing* data.

1. Pre-processing Data

Pre-processing data adalah sebuah teknik yang digunakan untuk mengambil informasi dari suatu data yang masih mentah. *Pre-processing* data dilakukan untuk menghilangkan noise yang terdapat pada data mentah, sehingga data yang nanti digunakan pada proses *data mining* adalah data dengan kualitas baik (Kasanah et al., 2019). Menurut Nofrizal (2021) tahapan *pre-processing* data terdiri dari:

a. *Data cleaning*

Tahapan *data cleaning* dilakukan dengan memodifikasi, mengubah, membersihkan data yang tidak akurat. Data yang diperoleh dari BMKG Provinsi Kalimantan Tengah terdapat beberapa data yang tidak terukur. Data tersebut dilakukan proses data cleaning dengan mencari nilai *average*. Tabel 1 adalah data yang telah melalui proses *data cleaning*.

Tabel 1. Hasil *Data Cleaning*

No.	Periode	Curah Hujan (mm)
1	09/07/2014	1,5
2	10/07/2014	0
3	11/07/2014	1,1
4	12/07/2014	0
5	13/07/2014	18
6	14/07/2014	1,5
7	15/07/2014	0
8	16/07/2014	1,5
9	17/07/2014	0
10	18/07/2014	0
⋮	⋮	⋮
4258	28/06/2023	0
4259	29/06/2023	0
4260	30/06/2023	0,2

b. *Data Transformation*

Tahapan *data transformation* dilakukan dengan mengubah data harian curah hujan yang telah dilakukan proses *data cleaning* menjadi data bulanan curah hujan. Tahapan *data transformation* dilakukan dengan mencari jumlah curah hujan pada setiap bulan. Table 2 adalah data bulanan curah hujan.

Tabel 2. Data Curah Hujan Bulanan

No.	Periode	Curah Hujan (mm)
1	07/2014	47
2	08/2014	66,5
3	09/2014	129,5
4	10/2014	123,3
5	11/2014	360,8
6	12/2014	695,4
7	01/2015	354,5
8	02/2015	545
9	03/2015	630
10	04/2015	495,1
⋮	⋮	⋮
138	04/2023	306,4
139	05/2023	134,33
140	06/2023	97,24

2. Pembagian Data

Proporsi pembagian data bersifat subjektif tergantung peneliti. Persentase *data training* lebih besar dibanding dengan *data testing*. Data testing yang dapat digunakan 10% sampai 30% dari data terakhir (Aisyah et al., 2021). Rasio skema pembagian data pada penelitian ini adalah rasio skema 70:30, 80:20, dan 90:10. Rasio skema pembagian data yang digunakan pada penerapan proses metode K-NN adalah rasio skema pembagian data terbaik.

3. Penerapan Proses Metode K-NN

Rasio skema pembagian data terbaik yang digunakan pada penerapan proses metode K-NN adalah rasio skema 70:30 dengan *data training* berjumlah 97 data terhitung dari Januari 2012 – Januari 2020 dan *data testing* berjumlah 41 data terhitung dari Februari 2020 – Juni 2023. *Data training* dikelompokkan menjadi 2 yaitu data input dan data target. Data input mempresentasikan 2 urutan data curah hujan sebelum data prediksi (x_1 dan x_2), sedangkan data target mempresentasikan data curah hujan yang diprediksi (y). Berikut adalah proses penerapan metode K-NN:

a. Penentuan nilai k. Tahapan penentuan nilai k tidak memiliki aturan khusus, namun pada penelitian ini nilai k terbaik adalah k =16.
b. Data yang digunakan pada penelitian ini adalah data bulanan curah hujan untuk memprediksi curah hujan. Data berikut adalah *data training* dan *testing* pada pembagian data 70:30. *Data training* terhitung sejak Januari 2012 – Januari 2020, sedangkan *data testing* terhitung sejak Februari 2020 – Juni

2023. *Data testing* yang dilampirkan pada penerapan proses hanya 1, namun perhitungan yang sama berlaku untuk semua *data testing*.

Tabel 3. *Data Training*

No.	x_1		x_2		y	
	Periode	**Actual**	**Periode**	**Actual**	**Periode**	**Actual**
1	11/2011	441,9	12/2011	388,9	01/2012	481,1
2	12/2011	388,9	01/2012	481,1	02/2012	265
3	01/2012	481,1	02/2012	265	03/2012	362,9
4	02/2012	265	03/2012	362,9	04/2012	310,3
5	03/2012	362,9	04/2012	310,3	05/2012	245,1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
95	09/2019	96,7	10/2019	253,5	11/2019	181,1
96	10/2019	253,5	11/2019	181,1	12/2019	430,8
97	11/2019	181,1	12/2019	430,8	01/2020	571,4

Tabel 4. *Data Testing*

No.	x_1		x_2		y	
	Periode	**Actual**	**Periode**	**Actual**	**Periode**	**Actual**
1	12/2019	430,8	01/2020	571,4	02/2020	?

c. Melakukan perhitungan nilai jarak antara *data training* dan *data testing*. Perhitungan nilai jarak terbaik pada penelitian ini adalah *Manhattan Distance*. Berikut perhitungan *Manhattan Distance*.

Tabel 1. Perhitungan *Manhattan Distance*

No.	x_1	x_2	Kuadrat jarak dengan data baru (430,8 dan 571,4)
1	441,9	388,9	| 430,8 – 441,9 | + | 571,4 – 388,9 | = 193,6
2	388,9	481,1	| 430,8 – 388,9 | + | 571,4 – 481,1 | = 132,2
3	481,1	265	| 430,8 – 481,1 | + | 571,4 – 265 | = 356,7
4	265	362,9	| 430,8 – 265 | + | 571,4 – 362,9 | = 374,3
5	362,9	310,3	| 430,8 – 441,9 | + | 571,4 – 310,3 | = 329
⋮	⋮	⋮	⋮
95	96,7	253,5	| 430,8 – 96,7 | + | 571,4 – 253,5 | = 652
96	253,5	181,1	| 430,8 – 253,5 | + | 571,4 – 181,1 | = 576,6
97	181,1	430,8	| 430,8 – 181,1 | + | 571,4 – 430,8 | = 390,3

d. Mengurutkan jarak dan menentukan tetangga terdekat berdasarkan dengan jarak terdekat ke-k (peringkat jarak minimum).
e. Menentukan kelompok data hasil perhitungan jarak berdasarkan dengan k tetangga terdekat. Karena k=16 maka diambil 16 jarak minimum.

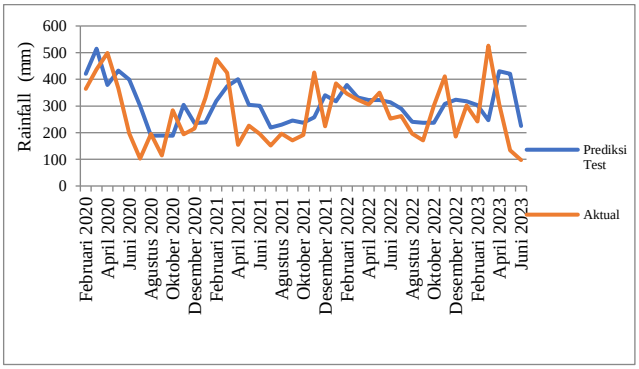
Tabel 2. Peringkat Jarak Minimum *Manhattan Distance*

Peringkat jarak minimum	Kuadrat jarak dengan data baru (430,8 dan 571,4)
1	| 430,8 – 424,9 | + | 571,4 – 527,5 | = 49,8
2	| 430,8 – 440,8 | + | 571,4 – 522,2 | = 59,2
3	| 430,8 – 408,9 | + | 571,4 – 533,8 | = 59,5
4	| 430,8 – 354,5 | + | 571,4 – 545 | = 102,7
5	| 430,8 – 533,8 | + | 571,4 – 543 | = 131,4
6	| 430,8 – 388,9 | + | 571,4 – 481,1 | = 132,2
7	| 430,8 – 339 | + | 571,4 – 617,2 | = 137,6
8	| 430,8 – 495,5 | + | 571,4 – 478,2 | = 157,9
9	| 430,8 – 545 | + | 571,4 – 630 | = 172,8
10	| 430,8 – 478,2 | + | 571,4 – 435,2 | = 183,6
11	| 430,8 – 313,9 | + | 571,4 – 639,1 | = 184,6
12	| 430,8 – 491,3 | + | 571,4 – 440,8 | = 191,1
13	| 430,8 – 441,9 | + | 571,4 – 388,9 | = 193,6
14	| 430,8 – 360,8 | + | 571,4 – 695,4 | = 194
15	| 430,8 – 270,1 | + | 571,4 – 610,6 | = 199,9
16	| 430,8 – 453,9 | + | 571,4 – 392,8 | = 201,7

f. Melakukan perhitungan rata-rata target dari k data teratas. Perhitungan rata-rata tersebut menghasilkan nilai hasil prediksi. Data yang digunakan pada perhitungan rata-rata adalah data aktual dari data teratas. Berikut perhitungan rata-rata pada *Manhattan Distance*.

$$y(t+1) = \frac{1}{k}\sum_{i=1}^{k} y_i$$
$$= \frac{k_1 + k_2 + k_3 + \cdots + k_{14} + k_{15} + k_{16}}{16}$$
$$= \frac{344{,}2 + 261{,}4 + 543 + \cdots + 354{,}4 + 347{,}4 + 98{,}5}{16}$$
$$= 411{,}25$$

g. Hasil perhitungan rata-rata pada tahap f merupakan nilai prediksi, maka dapat disimpulkan nilai prediksi curah hujan Februari 2020 berdasarkan perhitungan *Manhattan Distance* adalah 411,25 mm. Prediksi pada rentang Maret 2020 – Juni 2023 dilakukan dengan menggunakan perhitungan yang sama. Pada Gambar 2, memperlihatkan secara visual *trend* grafik *data testing* prediksi dengan *data testing* aktual, selanjutnya akan diuji dengan perhitungan RMSE.



Gambar 1. Prediksi Data *Testing*.

4. Evaluasi Menggunakan RMSE

Pelatihan dan evaluasi model dilakukan menggunakan RMSE seperti persamaan (4) pada masing-masing rasio skema pembagian data sehingga diperoleh banding hasil nilai RMSE terkecil. Skema pengujian dilakukan menggunakan RMSE dengan nilai k=16 pada perhitungan jarak *Manhattan Distance* dan *Euclidean Distance* dan berikut perhitungan manual RMSE.

Tabel 3. Perhitungan Manual Nilai RMSE

n	y_i	y_i'	$(y_i - y_i')^2$	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - yi')^2}$
1	411,25	364,7	46,55	
2	414,71	438	23,29	
3	366,61	499	132,39	109,03
⋮	⋮	⋮	⋮	
40	385,49	134,3	251,15	
41	211,26	97,24	114,02	

Evaluasi model menggunakan RMSE diperoleh perbandingan nilai RMSE pada setiap rasio skema pembagian data dengan perhitungan jarak *Manhattan Distance* dan *Euclidean Distance* dengan nilai k yang digunakan pada K-NN adalah 1 sampai 22. Perbandingan nilai RMSE tersebut dapat dilihat pada Tabel 8.

Tabel 4. Peringkat Jarak Minimum *Manhattan Distance*

nilai *k*	RMSE (Manhattan)			RMSE (Euclidean)		
	70:30	**80:20**	**90:10**	**70:30**	**80:20**	**90:10**
k = 1	180,5	181,6	201,6	193,8	185,5	201,2
k = 2	144,8	160,2	180,2	156,4	168,4	184,5
k = 3	128,0	141,5	153,0	128,2	147,4	161,8
k = 4	119,3	126,4	127,0	124,0	123,4	133,7
k = 5	120,6	126,4	133,1	123,2	124,3	132,8
k = 6	124,6	124,4	129,5	123,6	135,9	140,1
k = 7	123,6	125,3	132,1	121,8	133,9	146,7
k = 8	127,1	128,1	138,9	120,6	129,2	141,0
k = 9	126,7	127,6	140,1	124,5	124,7	133,1
k = 10	124,5	124,2	139,7	124,7	125,7	133,3
k = 11	127,7	130,4	136,9	118,5	128,0	134,5
k = 12	120,7	129,0	131,1	115,2	122,6	133,3
k = 13	118,6	127,5	132,6	114,9	118,9	130,7
k = 14	115,5	124,2	127,8	115,6	117,2	130,3
k = 15	112,9	122,9	127,9	116,6	118,6	127,1
k = 16	109,3	119,9	127,9	115,9	116,9	127,3
k = 17	110,1	118,1	125,6	115,1	115,6	123,9
k = 18	111,7	117,0	122,7	116,0	114,1	125,2
k = 19	112,4	113,5	120,5	113,6	116,6	123,4
k = 20	109,6	112,3	119,5	113,0	117,5	121,3
k = 21	111,1	113,7	118,5	112,6	116,4	121,6
k = 22	112,4	114,0	118,8	112,9	117,3	122,4

Tabel 8 menampilkan perbandingan nilai RMSE pada setiap rasio skema pembagian data terhadap *Manhattan Distance* dan *Euclidean Distance*. Berdasarkan Tabel 8 tersebut diperoleh bahwa performa terbaik adalah rasio skema pembagian data 70:30 pada perhitungan jarak *Manhattan Distance* dengan nilai k=16 dan nilai RMSE terkecil sebesar 109,3. Nilai RMSE terkecil mengindikasikan bahwa perhitungan jarak tersebut memiliki nilai error yang lebih rendah.

5. Hasil Prediksi Curah Hujan

Proses prediksi dilakukan pada rasio skema pembagian data 70:30 dengan nilai k=16 menggunakan *Manhattan Distance*. Hasil prediksi curah hujan selama 18 bulan terhitung dari Juli 2023 sampai Juni 2024 dalam interval bulanan dapat dilihat pada Tabel 9.

Tabel 5. Hasil Prediksi Curah Hujan

No	Periode	Curah Hujan (mm)
1	Juli 2023	199,94
2	Agustus 2023	202,23
3	September 2023	255,21
4	Oktober 2023	216,67
5	November 2023	208,36
6	Desember 2023	224,56
7	Januari 2024	241,2
8	Februari 2024	216,7
9	Maret 2024	217,66
10	April 2024	231,13
11	Mei 2024	218,51
12	Juni 2024	201,33

Berdasarkan Tabel 9 menunjukkan bahwa intensitas curah hujan terendah terjadi pada Juli 2023 dengan kapasitas curah hujan 199,94 mm dan intensitas hujan tertinggi terjadi pada September 2023 dengan kapasitas curah hujan 255,21 mm.

Merujuk pada Tabel 9 bahwa waktu untuk tanam padi di Kabupaten Pulang Pisau dapat dilakukan Bulan Juli, Agustus, dan November 2023, serta Juni 2024. Hal tersebut dikarenakan berdasarkan hasil prediksi curah hujan yang memenuhi syarat untuk penanaman padi yaitu dengan kapasitas curah hujan antara 100-200 mm/bulan dengan kebutuhan masa tanam padi yaitu membutuhkan waktu tanam selama 4 bulan.

KESIMPULAN

Berdasarkan penelitian ini, diperoleh kesimpulan bahwa hasil implementasi metode K-NN dalam memprediksi curah hujan di Kabupaten Pulang Pisau diperoleh nilai k terbaik adalah k = 16 pada rasio skema pembagian data 70:30 dengan perhitungan jarak terbaik yakni *Manhattan Distance* dan menghasilkan nilai error terkecil yaitu nilai RMSE sebesar 109,3. Hasil konversi dengan mempertimbangkan data *time series* curah hujan aktual menunjukkan bahwa Bulan Juli, Agustus, dan November 2023, serta Juni 2024 sebaiknya dilakukan proses tanam padi di Kabupaten Pulang Pisau. Hal ini cukup berbeda dengan kebiasaan yang dilakukan masyarakat untuk menanam padi pada sepanjang bulan November sampai Maret.

UCAPAN TERIMA KASIH

Terima kasih kepada Universitas Palangka Raya yang telah membiayai melalui Dana PNBP Anggaran Fakultas MIPA.

REFERENSI

Aisyah, S., Wahyuningsih, S., & Amijaya, F. (2021). Peramalan Jumlah Titik Panas Provinsi Kalimantan Timur Menggunakan Metode Radial Basis Function Neural Network. *Jambura Journal of Probability and Statistics*. Vol. 2(2), 64–

74. https://doi.org/10.34312/jjps.v2i2.10292

Arwansyah, & Arie, J. S. (2019). Implementasi Algoritma KNN Dalam Memprediksi Curah Hujan dan Temperatur Untuk Tanaman Padi. *Prosiding Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi*. Vol. 8(1), 11–20.

BMKG. (n.d.). *Data Online BMKG*. https://dataonline.bmkg.go.id/

Cahyati, U. (2019). Prediksi Curah Hujan Dengan Menggunakan Metode Jaringan Syaraf Tiruan (JST) Backpropagation Sebagai Pendukung Kalender Tanam Di Kabupaten Lamongan. *A Psicanalise Dos Contos de Fadas. Tradução Arlene Caetano*, 466.

Contena, E. (2022). *Optimasi Metode Jaringan Syaraf Tiruan Menggunakan Algoritma Genetika Dalam Prediksi Harga Cabai Di Samarinda*.

Evi Fitriana, M. (2021). Transmigran sebagai Modal Sosial dalam Pengembangan Food Estate di Kabupaten Pulang Pisau. *Jurnal Ilmiah Ilmu Sosial Dan Humaniora*. Vol. 7(1), 1–14.

Faidah, D. Y. (2022). Prediksi Curah Hujan Bulanan dengan North American Multi Model Ensemble. *PRISMA, Prosiding Seminar Nasional Matematika*. Vol. 5, 726–730. https://journal.unnes.ac.id/sju/index.php/prisma/article/view/54318%0A https://journal.unnes.ac.id/sju/index.php/prisma/article/download/54318/21103

Febriyanti, L. (2022). *Metode K-Nearest Neighbor Regression Dalam Memprediksi Indeks Pembangunan Manusia*. 1–23.

Kasanah, A. N., Muladi, M., & Pujianto, U. (2019). Penerapan Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Objektivitas Berita Online Menggunakan Algoritma KNN. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*. Vol. 3(2), 196–201. https://doi.org/10.29207/resti.v3i2.945

Laily, Y. H., Rakhmawati, F., Husein, I., Islam, U., Sumatera, N., & Medan, U. (2023). *PENERAPAN METODE FUZZY Penerapan Metode Fuzzy Time Series-Markov Chain Dalam Peramalan Curah Hujan Sebagai Jadwal Tanaman Padi*. Vol. 4(1).

Nikmatun, I. A., & Waspada, I. (2019). Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. *Jurnal SIMETRIS*. Vol. 10(2), 421–432.

Nisya, H. (2023). *Prediksi Curah Hujan Dari Data Satelit Himawari-8 Menggunakan Metode K-Nearest Neighbor (KNN)*. Vol. 10(1), 729–735.

Nofrizal, D. A. (2021). *Deteksi Fraud Pada Akun Wifi Universitas Islam Indonesia Dengan Metode K-Means*. https://dspace.uii.ac.id/handle/123456789/31785%0A https://dspace.uii.ac.id/bitstream/handle/123456789/31785/17523110

Rasyid, N. (2022). *Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Curah Hujan di Kota Bogor*. Vol. 1(1).

Rilwanu, F. N., Taufikurachman, H., Faris Huwaidi, D., Perangkat Lunak, R., & Daerah Cibiru, K. (2022). Penerapan Algoritma K-Nearest Neighbor untuk

Mendeteksi Diabetes Berbasis Web Application. *Journal of Software Engineering, Information and Communication Technology*. Vol. *3*(1), 145–152.

Yunita, S. & Nurahman. (2022). Metode Klasifikasi Data Mining Algoritma C4.5 Dan Part Untuk Prediksi Waktu Kelulusan Mahasiswa Di Universitas Darwan Ali. *Jurnal Informatika dan Teknologi Informasi: Informasi Interaktif*. Vol. 7(1), 1-7.

Zona kalteng. (2020). Petani di Desa Gadabung Gagal Tanam, Produksi Padi di Pulang Pisau Terancam Menurun. In *Online*. https://www.zonakalteng.co.id/2020/03/07/petani-di-desa-gadabung-gagal-tanam-produksi-padi-di-pulang-pisau-terancam-menurun/