

Prediksi Curah Hujan di Surabaya Berbasis Data BMKG Menggunakan Metode *K-Nearest Neighbor* (*K-NN*)

Alexander Daniel Kurniawan^{*1}, Nia Saurina², Teguh Pribadi Ikhsan³

^{1,2,3} Program Studi Informatika, Fakultas Teknik, Universitas Wijaya Kusuma Surabaya

Email: ¹adk014@mhs.uwks.ac.id, ²niasaurina@uwks.ac.id ³teguh@uwks.ac.id

^{*}Penulis Korespondensi

Abstrak

Curah hujan merupakan salah satu unsur cuaca yang berpengaruh terhadap berbagai aktivitas masyarakat, terutama di wilayah perkotaan seperti Surabaya. Perubahan intensitas hujan yang tidak menentu dapat menyebabkan genangan, banjir, serta gangguan aktivitas transportasi dan ekonomi. Oleh karena itu, diperlukan metode prediksi curah hujan yang mampu memberikan hasil cukup akurat berdasarkan data historis cuaca. Penelitian ini menerapkan algoritma *K-Nearest Neighbor* (*K-NN*) untuk memprediksi curah hujan harian menggunakan data BMKG periode 2021–2025 yang berasal dari Stasiun Meteorologi Perak I dan Stasiun Meteorologi Maritim Tanjung Perak. Variabel yang digunakan meliputi suhu, kelembapan udara, curah hujan, lama penyinaran matahari, dan kecepatan angin. Tahapan penelitian mencakup preprocessing data, normalisasi menggunakan *MinMaxScaler*, *feature engineering*, serta reduksi dimensi menggunakan *Principal Component Analysis* (*PCA*). Pengujian dilakukan menggunakan beberapa skenario pembagian data, variasi nilai *K*, dan fungsi jarak pada algoritma *K-NN*. Evaluasi model dilakukan menggunakan *MAE*, *RMSE*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Hasil penelitian menunjukkan bahwa kombinasi *feature engineering* dan *PCA* mampu meningkatkan performa model dalam memprediksi curah hujan harian berbasis data BMKG.

Kata kunci: Curah Hujan, *K-Nearest Neighbor*, *Machine Learning*, BMKG, Prediksi Cuaca, *PCA*

Abstract

Rainfall is one of the weather elements that significantly affects various community activities, especially in urban areas such as Surabaya. Unpredictable changes in rainfall intensity can cause flooding, transportation disruptions, and economic disturbances. Therefore, a rainfall prediction method capable of producing sufficiently accurate results based on historical weather data is required. This study applies the *K-Nearest Neighbor* (*K-NN*) algorithm to predict daily rainfall using BMKG data from 2020–2025 obtained from Perak I Meteorological Station and Tanjung Perak Maritime Meteorological Station. The variables used include temperature, humidity, rainfall, sunshine duration, and wind speed. The research stages include data preprocessing, normalization using *MinMaxScaler*, *feature engineering*, and dimensionality reduction using *Principal Component Analysis* (*PCA*). Testing was conducted using several data split scenarios, variations of *K* values, and distance metrics in the *K-NN* algorithm. Model evaluation was carried out using *MAE*, *RMSE*, *accuracy*, *precision*, *recall*, and *F1-score*. The results indicate that the combination of *feature engineering* and *PCA* improves the performance of the model in predicting daily rainfall based on BMKG data.

Keywords: Rainfall, *K-Nearest Neighbor*, *Machine Learning*, BMKG, Weather Prediction, *PCA*

I. PENDAHULUAN

Curah hujan merupakan salah satu unsur cuaca yang memiliki pengaruh besar terhadap berbagai aktivitas masyarakat, khususnya pada wilayah perkotaan seperti Kota Surabaya. Intensitas hujan yang berubah secara tidak menentu dapat menyebabkan berbagai permasalahan, seperti genangan air, banjir, gangguan transportasi, hingga hambatan aktivitas ekonomi dan transportasi [1]. Oleh karena itu, diperlukan suatu sistem prediksi curah hujan yang mampu memberikan informasi cuaca secara lebih cepat dan cukup akurat sebagai bentuk antisipasi terhadap kondisi cuaca yang akan terjadi.

Perkembangan *machine learning* memberikan alternatif pendekatan dalam melakukan prediksi cuaca berbasis data historis. Salah satu algoritma yang cukup sering digunakan adalah *K-Nearest Neighbor* (*K-NN*) karena memiliki implementasi yang sederhana dan mampu bekerja dengan baik pada data numerik [2], [3]. Algoritma *K-NN* melakukan prediksi berdasarkan kedekatan jarak antar data sehingga hasil prediksi dipengaruhi oleh pemilihan *distance metric* dan jumlah tetangga terdekat (nilai *K*) yang digunakan [4]. Beberapa penelitian sebelumnya menunjukkan bahwa metode *K-NN* mampu menghasilkan performa yang cukup baik pada proses klasifikasi maupun prediksi curah hujan [5], [6].

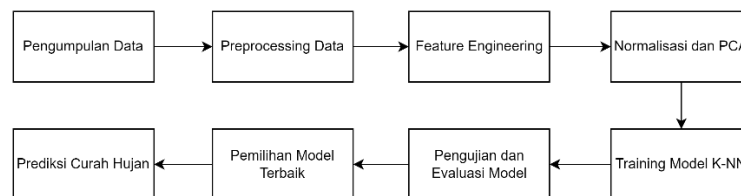
Data meteorologi yang digunakan dalam penelitian ini diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) berupa data curah hujan harian periode 2021–2025 dari Stasiun Meteorologi Perak I dan Stasiun Meteorologi Maritim Tanjung Perak [7]. Variabel yang digunakan meliputi suhu udara, kelembapan udara, curah hujan, lama penyinaran matahari, dan kecepatan angin. Namun, data cuaca umumnya memiliki pola yang kompleks dan fluktuatif sehingga diperlukan tahapan preprocessing dan pengolahan fitur agar model dapat mengenali pola data dengan lebih baik.

Pada penelitian ini diterapkan beberapa tahapan *preprocessing* seperti *data cleaning*, penanganan *missing value*, normalisasi data menggunakan *MinMaxScaler*, serta reduksi dimensi menggunakan *Principal Component Analysis (PCA)*. Selain itu dilakukan *feature engineering* berupa *lag feature*, *rolling window*, *delta feature*, dan fitur kalender untuk membantu model memahami pola temporal pada data cuaca [8]. Penggunaan *feature engineering* pada data *time series* dapat membantu model dalam mempelajari perubahan kondisi cuaca dari waktu ke waktu sehingga performa prediksi dapat meningkat.

Penelitian ini juga menggunakan beberapa skenario pembagian data train-test split, variasi nilai K , serta *distance metric* berupa *Euclidean*, *Manhattan*, dan *Minkowski* untuk memperoleh model terbaik [9]. Evaluasi model dilakukan menggunakan *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui performa model baik dari sisi regresi maupun klasifikasi [10]. Hasil dari penelitian ini diharapkan mampu memberikan model prediksi curah hujan yang cukup baik berdasarkan data historis BMKG di Kota Surabaya.

II. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang dimulai dari pengumpulan data BMKG, *preprocessing data*, *feature engineering*, normalisasi data dan *PCA*, *training model K-Nearest Neighbor (K-NN)*, pengujian serta evaluasi model, hingga pemilihan model terbaik untuk melakukan prediksi curah hujan berdasarkan data aktual BMKG.



Gambar 1 Alur Penelitian

2.1 Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data meteorologi dari BMKG yang berasal dari Stasiun Meteorologi Perak I dan Stasiun Meteorologi Maritim Tanjung Perak periode 2021–2025. Data yang digunakan meliputi curah hujan, suhu udara, kelembapan, kecepatan angin, dan parameter cuaca lainnya yang digunakan sebagai fitur prediksi.

2.2 Preprocessing Data

Tahap *preprocessing* dilakukan untuk mempersiapkan data sebelum digunakan pada proses machine learning. Proses ini meliputi pembersihan data, penanganan *missing value*, penggabungan data antar stasiun, serta penyesuaian nama dan format kolom agar data lebih konsisten dan siap digunakan pada tahap berikutnya.

2.3 Feature Engineering

Pada tahap *feature engineering* dilakukan penambahan beberapa fitur baru seperti *lag feature*, *rolling feature*, *difference feature*, dan *calendar feature* untuk membantu model memahami pola historis curah hujan dan perubahan kondisi cuaca dari waktu ke waktu.

2.4 Normalisasi dan PCA

Data yang telah diproses kemudian dinormalisasi menggunakan *MinMaxScaler* agar seluruh fitur memiliki rentang nilai yang seimbang. Setelah itu dilakukan reduksi dimensi menggunakan *Principal Component Analysis (PCA)* untuk mengurangi kompleksitas fitur dan meningkatkan efisiensi proses training model.

2.5 Training Model K-NN

Tahap training dilakukan menggunakan algoritma *K-Nearest Neighbor (K-NN)* dengan beberapa kombinasi *distance metric* yaitu *Euclidean*, *Manhattan*, dan *Minkowski* serta pengujian beberapa nilai *K*. Selain itu diterapkan *cross validation* untuk memperoleh hasil evaluasi model yang lebih stabil.

2.6 Pengujian dan Evaluasi Model

Pada tahap ini dilakukan pengujian performa model menggunakan evaluasi regresi berupa *Mean Absolute Error (MAE)* dan *Root Mean Squared Error (RMSE)*, serta evaluasi klasifikasi berupa *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

2.7 Pemilihan Model Terbaik

Model terbaik dipilih berdasarkan hasil evaluasi yang diperoleh dari seluruh kombinasi parameter dan skenario pengujian. Pemilihan model dilakukan dengan mempertimbangkan nilai error terkecil dan performa klasifikasi terbaik.

2.8 Prediksi Curah Hujan

Model terbaik yang diperoleh kemudian digunakan untuk melakukan prediksi curah hujan periode Januari–April 2026. Hasil prediksi tersebut selanjutnya dibandingkan dengan data aktual BMKG untuk mengetahui kemampuan model dalam melakukan prediksi pada data baru.

III. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Skenario dan Parameter Model

Pada tahap ini dilakukan pengujian algoritma *K-Nearest Neighbor (K-NN)* menggunakan beberapa kombinasi parameter seperti nilai *K*, metode perhitungan jarak, serta skenario pembagian data training dan testing. Pengujian dilakukan menggunakan data curah hujan BMKG Surabaya periode 2021–2025 yang telah melalui proses *preprocessing*, normalisasi, *feature engineering*, dan reduksi dimensi menggunakan *PCA*. Tujuan dari pengujian ini adalah untuk mengetahui konfigurasi model yang mampu memberikan hasil prediksi paling optimal.

Ratio	Metric	Best_K	MAE	TEST_RMSE	Best_K	ACC	TEST_Accuracy (%)	TEST_Precision	TEST_Recall	TEST_F1	PCA_Component
0	0.60 euclidean	9	5.46	12.22	7	64.12	0.70	0.64	0.64	11	1
1	0.60 manhattan	9	5.46	12.00	3	63.00	0.67	0.63	0.64	11	
2	0.60 minkowski	9	5.46	12.22	7	64.12	0.70	0.64	0.64	11	
3	0.70 euclidean	9	5.52	11.18	3	70.50	0.74	0.70	0.71	10	2
4	0.70 manhattan	7	5.38	11.29	3	70.13	0.74	0.70	0.71	10	
5	0.70 minkowski	9	5.52	11.18	3	70.50	0.74	0.70	0.71	10	
6	0.80 euclidean	5	6.44	12.05	7	62.22	0.71	0.62	0.62	10	3
7	0.80 manhattan	7	6.32	12.22	7	61.39	0.69	0.61	0.60	10	
8	0.80 minkowski	5	6.44	12.05	7	62.22	0.71	0.62	0.62	10	
9	0.90 euclidean	3	5.54	12.23	7	68.89	0.75	0.69	0.70	10	4
10	0.90 manhattan	3	5.18	11.37	3	69.44	0.74	0.69	0.70	10	
11	0.90 minkowski	3	5.54	12.23	7	68.89	0.75	0.69	0.70	10	

Gambar 2 Hasil Terbaik Skenario Per Metric

Hasil pengujian menunjukkan bahwa setiap kombinasi parameter memberikan performa yang berbeda terhadap nilai prediksi curah hujan. Beberapa konfigurasi menghasilkan nilai *error* yang rendah dan tingkat akurasi yang tinggi, sedangkan konfigurasi lainnya menunjukkan performa yang kurang stabil. Oleh karena itu, dilakukan perbandingan hasil evaluasi untuk menentukan model terbaik yang akan digunakan pada tahap prediksi data aktual.

3.2 Hasil Evaluasi Model Terbaik Berdasarkan MAE

Evaluasi model dilakukan menggunakan metrik numerik berupa *Mean Absolute Error (MAE)* dan *Root Mean Squared Error (RMSE)* untuk mengukur tingkat kesalahan prediksi nilai curah hujan. Selain itu, digunakan juga evaluasi klasifikasi seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix* untuk mengetahui kemampuan model dalam mengklasifikasikan kategori intensitas hujan berdasarkan standar BMKG.

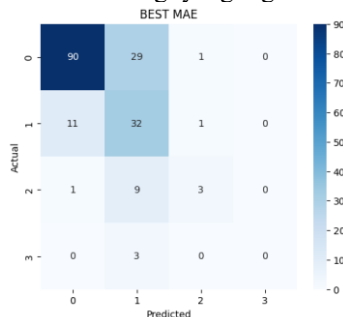
```
=== BEST BY MAE ===
ratio          0.90
metric         manhattan
k              3
CV_MAE         6.10
CV_RMSE        12.20
CV_Accuracy     0.61
CV_Precision    0.65
CV_Recall       0.61
CV_F1          0.61
TEST_MAE        5.18
TEST_RMSE       11.37
TEST_Accuracy   0.69
TEST_Precision  0.74
TEST_Recall     0.69
TEST_F1        0.70
PCA_Component   10
Name: 37, dtype: object
```

Gambar 3 Hasil Evaluasi Terbaik berdasarkan *MAE*

Berdasarkan hasil evaluasi, model terbaik diperoleh dari kombinasi parameter yang mampu menghasilkan nilai *MAE* dan *RMSE* paling kecil serta akurasi klasifikasi paling tinggi. Hasil confusion matrix menunjukkan bahwa model mampu mengenali kategori hujan ringan, sedang, dan lebat dengan cukup baik meskipun masih terdapat beberapa kesalahan klasifikasi pada kondisi curah hujan tertentu.

3.3 Analisis Confusion Matrix Model Terbaik berdasarkan *MAE*

Confusion matrix digunakan untuk mengetahui kemampuan model dalam mengklasifikasikan kategori curah hujan berdasarkan hasil prediksi terbaik yang diperoleh dari nilai *MAE* terkecil. Dari hasil *confusion matrix* terlihat bahwa sebagian besar data berhasil diklasifikasikan dengan benar, terutama pada kategori hujan ringan dan sedang. Hal ini menunjukkan bahwa model *K-NN* mampu mengenali pola data curah hujan berdasarkan kemiripan fitur meteorologi yang digunakan pada proses training.

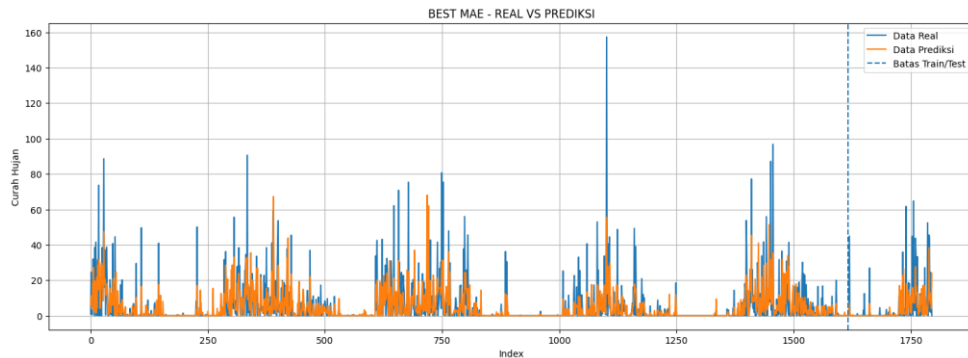


Gambar 4 *Confusion Matrix* Model Terbaik Berdasarkan *MAE*

Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi pada kategori hujan dengan intensitas lebih tinggi. Kesalahan tersebut dipengaruhi oleh pola data curah hujan yang fluktuatif serta distribusi data yang tidak seimbang antar kategori hujan. Namun secara keseluruhan, *confusion matrix* menunjukkan bahwa model memiliki performa klasifikasi yang cukup baik dalam memprediksi kategori curah hujan harian di Surabaya.

3.4 Analisis Grafik Model Terbaik berdasarkan *MAE*

Grafik perbandingan antara data aktual dan hasil prediksi digunakan untuk melihat sejauh mana model mampu mengikuti pola perubahan curah hujan harian. Berdasarkan grafik yang diperoleh, hasil prediksi model terlihat cukup mengikuti pola data aktual, terutama pada kondisi curah hujan rendah hingga sedang. Hal ini menunjukkan bahwa algoritma *K-NN* mampu mempelajari pola historis data cuaca dengan cukup baik setelah dilakukan preprocessing dan feature engineering.



Gambar 5 Grafik Model Terbaik Berdasarkan *MAE*

Namun pada beberapa periode tertentu masih terdapat selisih antara nilai aktual dan nilai prediksi, khususnya pada kondisi hujan dengan intensitas tinggi. Perbedaan tersebut disebabkan oleh karakteristik data cuaca yang sangat dinamis dan sulit diprediksi secara sempurna. Meskipun demikian, secara umum grafik menunjukkan bahwa model mampu memberikan hasil prediksi yang stabil dan cukup representatif terhadap kondisi curah hujan aktual.

3.5 Hasil Evaluasi Model Terbaik Berdasarkan *Accuracy*

Selain berdasarkan nilai *MAE*, evaluasi model juga dilakukan menggunakan metrik *accuracy* untuk mengetahui kemampuan model dalam mengklasifikasikan kategori curah hujan secara keseluruhan. Hasil pengujian menunjukkan bahwa kombinasi parameter tertentu mampu menghasilkan tingkat *accuracy* yang lebih tinggi dibandingkan kombinasi lainnya. Model terbaik berdasarkan *accuracy* diperoleh dari kombinasi nilai *K* dan *distance metric* yang mampu memberikan keseimbangan antara hasil prediksi numerik dan klasifikasi kategori hujan.

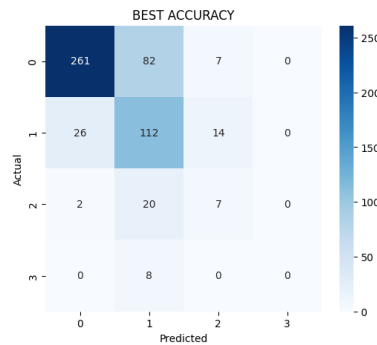
```
=== BEST BY ACCURACY ===
ratio      0.70
metric     euclidean
k          3
CV_MAE     6.10
CV_RMSE    11.84
CV_Accuracy 0.62
CV_Precision 0.65
CV_Recall  0.62
CV_F1      0.62
TEST_MAE   5.62
TEST_RMSE  12.23
TEST_Accuracy 0.71
TEST_Precision 0.74
TEST_Recall 0.71
TEST_F1    0.71
PCA_Component 10
Name: 12, dtype: object
```

Gambar 6 Hasil Evaluasi Terbaik Berdasarkan *Accuracy*

Nilai *accuracy* yang diperoleh menunjukkan bahwa model mampu mengenali sebagian besar kategori hujan dengan cukup baik. Selain itu, nilai *precision*, *recall*, dan *F1-score* juga menunjukkan performa yang relatif stabil pada beberapa kategori hujan. Dengan demikian, model terbaik berdasarkan *accuracy* dapat digunakan sebagai alternatif model prediksi curah hujan harian berbasis data BMKG di Kota Surabaya.

3.6 Analisis *Confusion Matrix* Model Terbaik berdasarkan *Accuracy*

Confusion matrix pada model terbaik berdasarkan *accuracy* digunakan untuk melihat distribusi hasil klasifikasi antara data aktual dan data prediksi pada setiap kategori hujan. Berdasarkan hasil *confusion matrix*, sebagian besar prediksi berada pada diagonal utama yang menunjukkan bahwa model berhasil melakukan klasifikasi dengan benar pada mayoritas data pengujian. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik terhadap data baru.

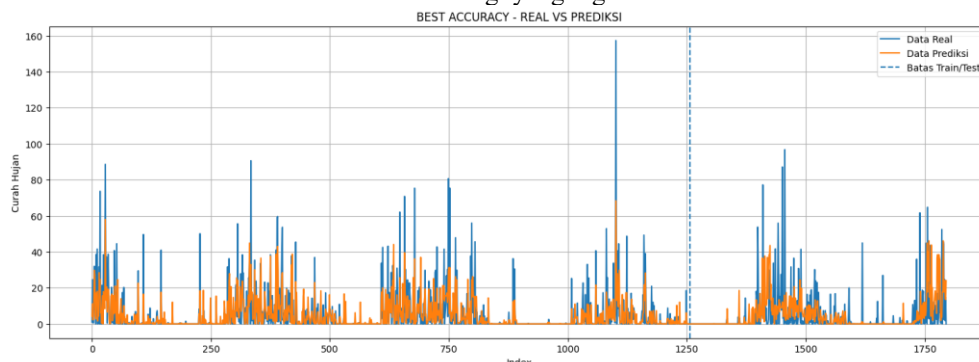


Gambar 7 Confusion Matrix Model Terbaik Berdasarkan Accuracy

Walaupun demikian, masih terdapat beberapa data yang salah diklasifikasikan ke kategori lain, terutama pada kategori hujan sedang dan hujan lebat yang memiliki pola data lebih kompleks. Kesalahan tersebut dipengaruhi oleh kemiripan karakteristik antar kategori hujan sehingga model mengalami kesulitan dalam membedakan beberapa kondisi tertentu. Namun secara keseluruhan, *confusion matrix* menunjukkan bahwa model memiliki tingkat klasifikasi yang cukup baik dan stabil.

3.7 Analisis Grafik Matrix Model Terbaik berdasarkan Accuracy

Grafik hasil prediksi model terbaik berdasarkan *accuracy* menunjukkan perbandingan antara data curah hujan aktual dan hasil prediksi model pada data pengujian. Dari grafik terlihat bahwa pola prediksi model mampu mengikuti tren perubahan curah hujan harian dengan cukup baik, terutama pada periode dengan intensitas hujan rendah hingga menengah. Hal ini menunjukkan bahwa model mampu mengenali pola historis data cuaca berdasarkan fitur meteorologi yang digunakan.

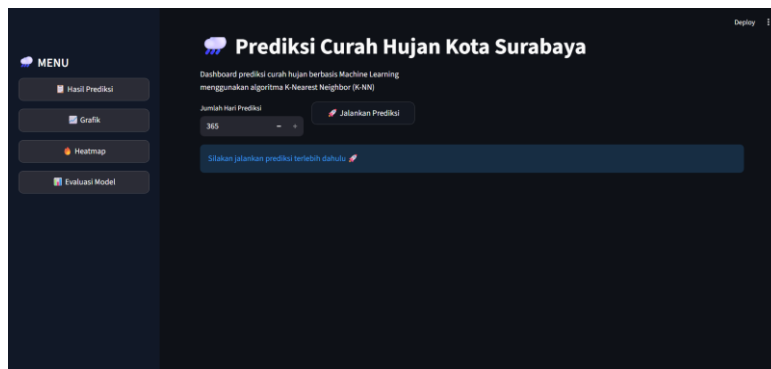


Gambar 8 Grafik Model Terbaik Berdasarkan Accuracy

Pada beberapa titik tertentu masih terlihat perbedaan antara nilai aktual dan nilai prediksi, khususnya ketika terjadi lonjakan curah hujan yang cukup tinggi. Kondisi tersebut menunjukkan bahwa data cuaca memiliki sifat yang sangat fluktuatif sehingga sulit diprediksi secara sempurna. Namun secara umum, grafik menunjukkan bahwa model *K-NN* mampu memberikan hasil prediksi yang cukup stabil dan dapat digunakan untuk membantu analisis prediksi curah hujan harian di Kota Surabaya.

3.8 Tampilan Halaman Utama Dashboard

Gambar 9 menunjukkan implementasi antarmuka aplikasi prediksi curah hujan berbasis *web* yang dikembangkan menggunakan *framework Streamlit*. Aplikasi ini dirancang untuk mempermudah proses prediksi curah hujan Kota Surabaya menggunakan model *K-Nearest Neighbor (K-NN)* yang telah melalui tahapan *preprocessing data*, *feature engineering*, normalisasi menggunakan *MinMaxScaler*, serta reduksi dimensi menggunakan *Principal Component Analysis (PCA)*. Pada halaman utama ditampilkan judul aplikasi beserta deskripsi singkat mengenai metode yang digunakan. Selain itu, pengguna dapat menentukan jumlah hari prediksi yang diinginkan melalui input "Jumlah Hari Prediksi", kemudian menjalankan proses forecasting menggunakan tombol "Jalankan Prediksi". Proses prediksi dilakukan secara *autoregressive forecasting*, yaitu hasil prediksi curah hujan pada hari sebelumnya akan digunakan kembali sebagai input untuk memprediksi curah hujan pada hari berikutnya.



Gambar 9 Tampilan Halaman Utama Dashboard

Antarmuka sistem dirancang dalam bentuk *dashboard* interaktif dengan tampilan *dark mode* agar lebih modern dan nyaman digunakan. Pada bagian *sidebar* tersedia beberapa menu navigasi seperti hasil prediksi, grafik, *heatmap*, dan evaluasi model untuk mempermudah pengguna dalam mengakses informasi yang dihasilkan. Halaman hasil prediksi digunakan untuk menampilkan data *forecasting* curah hujan beserta kategori intensitas hujan, sedangkan halaman grafik digunakan untuk memvisualisasikan pola hasil prediksi dalam bentuk *line chart*. Selain itu, halaman *heatmap* digunakan untuk menampilkan distribusi intensitas curah hujan berdasarkan periode waktu tertentu, sementara halaman evaluasi model digunakan untuk menampilkan performa model terbaik berdasarkan nilai *MAE*, *RMSE*, *accuracy*, *precision*, *recall*, dan *F1-score*. Dengan adanya sistem berbasis *web* ini, proses analisis dan pemantauan prediksi curah hujan dapat dilakukan secara lebih mudah, informatif, dan interaktif.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, algoritma *K-Nearest Neighbor (K-NN)* berhasil diterapkan untuk prediksi curah hujan harian di Kota Surabaya menggunakan data meteorologi BMKG periode 2021–2025. Penelitian dilakukan melalui beberapa tahapan, yaitu *preprocessing data*, *feature engineering*, normalisasi menggunakan *MinMaxScaler*, serta reduksi dimensi menggunakan *Principal Component Analysis (PCA)*. Pengujian model dilakukan menggunakan beberapa skenario pembagian data, variasi nilai *K*, serta *distance metric* berupa *Euclidean*, *Manhattan*, dan *Minkowski* untuk memperoleh konfigurasi model terbaik. Hasil evaluasi menunjukkan bahwa kombinasi *feature engineering* dan *PCA* mampu membantu meningkatkan performa model dalam mengenali pola data curah hujan yang kompleks dan fluktuatif.

Evaluasi model dilakukan menggunakan metrik *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, *accuracy*, *precision*, *recall*, dan *F1-score*. Berdasarkan hasil pengujian, beberapa kombinasi parameter mampu menghasilkan performa yang cukup baik dalam proses prediksi maupun klasifikasi kategori curah hujan. Selain itu, penggunaan *cross validation* membantu menghasilkan evaluasi model yang lebih stabil dan mengurangi bias pada proses pengujian. Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma *K-Nearest Neighbor (K-NN)* dapat digunakan sebagai salah satu metode *machine learning* yang cukup efektif untuk prediksi curah hujan berbasis data BMKG di Kota Surabaya serta dapat dikembangkan lebih lanjut pada penelitian berikutnya.

REFERENSI

- [1] I. D. G. L. Maheswara dan A. H. Al'aziz, "Perbandingan Model Machine Learning pada Klasifikasi Curah Hujan di Bogor," *Jurnal INTI*, 2025.
- [2] M. A. Ramadhan, F. T. Anggraeny dan C. A. Putra, "Klasifikasi Curah Hujan Harian Menggunakan Metode K-Nearest Neighbor (KNN)," *JATI(Jurnal Mahasiswa Teknik Informatika)*, 2024.
- [3] I. Ismail, M. A. Wardana dan L. Sarwana, "Implementasi Metode K-Nearest Neighbors Untuk Prediksi Cuaca Dengan Konsep Data Mining," *Jurnal RISTER*, pp. 59-66, 2024.
- [4] M. A. Aprihartha, T. N. Alam dan M. Husniyadi, "Perbandingan Metrik Euclidean dan Metrik Manhattan untuk K-Nearest Neighbors dalam Klasifikasi Kismis," *Jurnal Ilmu Komputer dan Informatika (JIKI)*, 2024.

-
- [5] D. M. Nanda, T. H. Pudjiantoro dan P. N. Sabrina, “Metode K-Nearest Neighbor (KNN) Dalam Memprediksi Curah Hujan di Kota Bandung,” *Prosiding Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika (SNESTIK)*, 2022.
 - [6] A. S. Agung, A. Fauzi, A. A. N. Risal dan F. Adiba, “Implementasi Teknik Data Mining terhadap Klasifikasi Data Prediksi Curah Hujan BMKG di Sulawesi Selatan.,” *Jurnal Tekno Insentif*, 2023.
 - [7] BMKG, “Layanan Data dan Informasi Meteorologi,” BMKG (Portal Data Resmi), 2024.
 - [8] G. A. RAKHMAT dan W. MUTOHAR, “Prakiraan Hujan menggunakan Metode Random Forest dan Cross Validation,” *MIND JOURNAL (Multimedia artificial Intelligence Networking Database)*, 2023.
 - [9] A. Setiawan, “Perbandingan Penggunaan Jarak Manhattan, Jarak Euclid, dan Jarak Minkowski dalam Klasifikasi Menggunakan Metode KNN pada Data Iris,” *Jurnal Sains dan Edukasi Sains*, 2022.
 - [10] T. O. Hodson, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” *Geoscientific Model Development*, 2022.