

Model Prediksi Risiko Banjir Menggunakan Algoritma Machine Learning: Kajian Literatur dan Aplikasi Metode Decision Tree

Abiyanfaras Danuyasa *¹
Hafidza Dfariz Mujizat ²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Indonesia
*e-mail: abiyanfaras@mhs.pelitabangsa.ac.id¹, hafidza12367@mhs.pelitabangsa.ac.id²,

Abstrak

Dengan menggunakan metode penelitian literatur, penelitian ini menyelidiki penerapan algoritma pembelajaran mesin, khususnya Decision Tree, dalam memprediksi risiko banjir. Banjir adalah bencana alam yang sering terjadi di negara tropis seperti Indonesia, terutama di daerah dataran rendah seperti Kota Pontianak. Penelitian ini melihat keakuratan, variabel input, dan performa algoritma pemodelan banjir melalui tinjauan sepuluh jurnal ilmiah. Selain itu, penelitian ini mengembangkan kerangka konseptual untuk sistem prediksi banjir berbasis Decision Tree yang menguji data klimatologi harian. Hasilnya menunjukkan bahwa faktor-faktor seperti kelembapan udara, curah hujan, dan penyinaran matahari sangat memengaruhi klasifikasi kondisi banjir. Meskipun model menunjukkan akurasi 87,7%, kesalahan prediksi minoritas tetap ada. Penelitian ini menunjukkan bahwa penggabungan pembelajaran mesin ke dalam sistem peringatan dini bencana sangat menjanjikan untuk pengambilan keputusan yang cepat dan fleksibel di wilayah yang rentan terhadap banjir.

Kata kunci: algoritma decision tree, banjir, klasifikasi, machine learning, prediksi

Abstract

Using a literature review method, this study investigates the application of machine learning algorithms, particularly Decision Tree, in predicting flood risk. Floods are natural disasters that frequently occur in tropical countries such as Indonesia, especially in lowland areas like the city of Pontianak. This research examines the accuracy, input variables, and performance of flood modeling algorithms through a review of ten scientific journals. In addition, the study develops a conceptual framework for a Decision Tree-based flood prediction system that evaluates daily climatological data. The results show that factors such as air humidity, rainfall, and sunlight duration significantly influence flood condition classification. Although the model demonstrates an accuracy of 87.7%, misclassification of minority classes still occurs. This research suggests that integrating machine learning into disaster early warning systems is highly promising for fast and flexible decision-making in flood-prone areas.

Keywords: decision tree algorithm, flood, classification, machine learning, prediction

PENDAHULUAN

Salah satu bencana alam yang paling sering terjadi di dunia adalah banjir. Ini menyebabkan kerugian ekonomi, sosial, dan kerusakan lingkungan yang signifikan. Laporan yang dikeluarkan oleh World Bank dan World Meteorological Organization (WMO) menunjukkan bahwa selama dua puluh tahun terakhir, banjir telah mengakibatkan kerugian ekonomi sebesar ratusan miliar dolar setiap tahun. Karena iklim tropisnya, curah hujan tinggi, dan pertumbuhan kota yang pesat tanpa perencanaan infrastruktur yang memadai, Asia Tenggara, termasuk Indonesia, sangat rentan terhadap bencana ini. Menurut data yang dikumpulkan oleh Badan Nasional Penanggulangan Bencana (BNPB), banjir adalah bencana yang paling sering terjadi di Indonesia. Alih fungsi lahan, deforestasi, penurunan kapasitas resapan tanah, dan sistem drainase yang tidak efisien di daerah perkotaan adalah beberapa penyebab utama tingginya frekuensi banjir di tanah air.

Kota Pontianak, yang terletak di garis khatulistiwa dan dilintasi oleh Sungai Kapuas, adalah salah satu daerah di Indonesia yang memiliki ketinggian rendah (sekitar 0,1 hingga 1,5 meter di atas permukaan laut). Ini merupakan contoh nyata daerah yang sangat rentan terhadap banjir, terutama selama musim hujan (Fitrah et al., 2025). Pontianak harus menjadi pusat

pengembangan sistem mitigasi dan adaptasi banjir karena kondisi geografisnya, urbanisasi yang pesat, dan ketergantungan terhadap sungai besar. Di daerah yang rawan banjir, tidak hanya terjadi kerusakan fisik, seperti kerusakan infrastruktur, pemukiman, dan fasilitas umum, tetapi juga memiliki konsekuensi jangka panjang, seperti masalah kesehatan masyarakat, gangguan pendidikan, masalah mata pencaharian, dan peningkatan risiko kemiskinan.

Oleh karena itu, dibutuhkan pendekatan terpadu yang tidak hanya reaktif tetapi juga prediktif dan preventif. Ini termasuk membangun sistem peringatan dini yang adaptif dan presisi tinggi dengan menggunakan teknologi informasi dan data spasial. Pendekatan data-driven berbasis kecerdasan buatan (Artificial Intelligence/AI), khususnya machine learning (ML), semakin banyak digunakan dalam sistem prediksi bencana sebagai akibat dari peningkatan teknologi digital dan komputasi awan. ML memungkinkan pengolahan data historis dan real-time untuk mengidentifikasi pola kompleks yang tidak dapat ditemukan dengan metode analisis tradisional. Dalam hal prediksi banjir, teknologi ini sangat bermanfaat karena dapat memodelkan hubungan non-linear antara berbagai parameter meteorologi, termasuk curah hujan, tinggi muka air, kelembaban udara, topografi, tata guna lahan, dan variabel lainnya.

Dalam penelitian tentang prediksi banjir, berbagai algoritma pembelajaran mesin telah digunakan. Ini termasuk Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, k-Nearest Neighbour (kNN), dan Neural Network. Studi yang dilakukan oleh Mosavi et al. (2018) menemukan bahwa pendekatan kecerdasan buatan lebih baik daripada model statistik konvensional dalam memprediksi kejadian banjir. Studi yang dilakukan oleh Chen et al. (2021) menemukan bahwa model seperti Gradient Boosting Decision Tree (GBDT), eXtreme Gradient Boosting (XGBoost), dan Convolutional Neural Network (CNN) sangat baik dalam menggambarkan risiko banjir di wilayah metropolitan. Namun, penelitian yang dilakukan oleh Razali et al. (2020) menemukan bahwa memprediksi risiko banjir di Malaysia dengan akurasi 99.92% dapat dicapai dengan menggunakan Decision Tree dan teknik SMOTE untuk penyeimbangan data.

Berdasarkan latar belakang ini, artikel ini ditulis untuk merangkum berbagai pendekatan baru untuk prediksi banjir berbasis ML dengan fokus pada algoritma Decision Tree. Pendekatan ini dipilih karena memiliki beberapa keuntungan, termasuk interpretabilitas, efisiensi, dan kemudahan implementasi. Selain itu, artikel ini mengusulkan sebuah kerangka kerja awal untuk sistem prediksi banjir yang dapat digunakan untuk daerah seperti Kota Pontianak. Akibatnya, diharapkan bahwa penelitian ini akan memberikan kontribusi ilmiah dan praktis untuk proses pengembangan sistem prediksi dan mitigasi bencana banjir yang adaptif, presisi tinggi, dan berbasis kecerdasan buatan.

METODE

Studi ini menggunakan kajian literatur sistematis terhadap berbagai publikasi ilmiah tentang prediksi banjir berbasis machine learning. Studi kasus dari China (Wei et al., 2019), Bangladesh (Syeed et al., 2021), Malaysia (Razali et al., 2020), dan Indonesia (Fitrah et al., 2025) termasuk dalam kategori artikel yang dikaji. Akurasi model, variabel input, dan jenis algoritma adalah semua parameter yang dievaluasi.

Proses analisis termasuk:

1. Identifikasi sumber-sumber utama dari jurnal yang terdaftar:

Tahap ini mencakup proses mencari dan memilih literatur yang relevan menggunakan database akademik seperti ScienceDirect, IEEE Xplore, SpringerLink, dan Google Scholar. "Prediction of floods", "machine learning", "decision tree", dan "mapping risiko banjir" adalah kata kunci yang digunakan. Artikel yang terpilih memiliki data eksperimen atau validasi model yang jelas dan membahas penggunaan algoritma klasifikasi dalam prediksi banjir.

2. Ekstrak data dari model prediksi, teknik, keakuratan, dan tempat penelitian:

Pada langkah ini, setiap artikel yang dipilih secara menyeluruh dianalisis. Analisis ini mencakup informasi tentang berbagai jenis algoritma pengajaran mesin yang digunakan; variabel input seperti curah hujan, elevasi, dan penggunaan lahan; ukuran dataset; dan nilai akurasi atau metrik evaluasi lainnya. Untuk memudahkan interpretasi dan analisis lebih lanjut, data ini disusun dalam bentuk tabel perbandingan studi.

3. Perbandingan kinerja algoritma pengelompokan:

Setelah data dikumpulkan, analisis perbandingan dilakukan untuk membandingkan kinerja masing-masing algoritma. Faktor-faktor seperti akurasi, sensitivitas, presisi, dan keandalan model terhadap data tidak seimbang dibandingkan. Selain itu, evaluasi mempertimbangkan manfaat dan kekurangan masing-masing algoritma dalam konteks topografi yang berbeda. Fokus utama penelitian ini adalah membandingkan Decision Tree dengan algoritma lain untuk menilai seberapa efektif dan konsisten mereka.

4. Perancangan sistem prediksi banjir berbasis Decision Tree untuk studi kasus lokal:

Berdasarkan analisis literatur, dirumuskan kerangka kerja konseptual yang memanfaatkan algoritma Decision Tree dalam sistem prediksi banjir di Kota Pontianak. Rangkaian ini mencakup langkah-langkah: (a) pengumpulan serta integrasi data curah hujan historis, tinggi muka air, dan data geografi; (b) pra-pemrosesan data mencakup normalisasi dan pemilihan fitur; (c) pelatihan model Pohon Keputusan; (d) penilaian model dengan memanfaatkan confusion matrix dan metrik lain; serta (e) penerapan dalam bentuk sistem peringatan dini berbasis web atau GIS. Rangka ini diajukan sebagai langkah pertama untuk menuju sistem prediksi banjir yang tepat dan berbasis teknologi di tingkat lokal.

HASIL DAN PEMBAHASAN

Implementasi Kode dan Pra - Pemrosesan Data

Pustaka pandas digunakan untuk mengimpor data dari berkas data_finish.csv untuk memulai analisis. Untuk analisis, atribut numerik seperti Tn, Tx, Tavg, RH_avg, RR, ss, ff_x, ddd_x, dan ff_avg telah dipilih. Semua atribut ini berkaitan dengan cuaca dan kemungkinan banjir. Jika tidak ada hujan, nilai yang hilang pada kolom curah hujan (RR) diatur menjadi nol. Untuk menjaga kestabilan distribusi data, nilai yang hilang pada fitur lain dipenuhi dengan nilai median. Agar tidak mengganggu proses pelatihan model, kolom non-numerik seperti date, station_id, station_name, region_name, dan ddd_car juga dihilangkan.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
import numpy as np

# Load dataset
df = pd.read_csv('data_finish.csv')

# Pra-pemrosesan:
df['RR'] = df['RR'].fillna(0) # curah hujan kosong dianggap nol
num_cols = ['Tn', 'Tx', 'Tavg', 'RH_avg', 'ss', 'ff_x', 'ddd_x', 'ff_avg']
for col in num_cols:
    df[col] = df[col].fillna(df[col].median())

# Hapus kolom non-numerik
df = df.drop(columns=['date', 'station_id', 'station_name', 'region_name', 'ddd_car'])

# Fitur dan label
X = df.drop(columns='flood')
y = df['flood']

# Bagi data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Gambar 1. Kode Pra - Pemrosesan Data

Setelah data dibersihkan, fitur dan label dipisahkan. Sementara variabel X menyimpan semua fitur prediktor, label target y menyimpan status banjir. Kemudian, fungsi train_test_split digunakan untuk membagi data menjadi data pelatihan dan data pengujian dengan perbandingan 80:20. Langkah ini sangat penting untuk memastikan bahwa model dapat dilatih dan dievaluasi dengan benar pada data yang belum pernah dilihat sebelumnya, dan untuk mengevaluasi seberapa baik model dapat menggeneralisasi pada data baru.

Confusion Matrix dan Feature Importance

Untuk menilai kinerja dan interpretabilitas model, dua jenis visualisasi utama digunakan untuk melakukan evaluasi. Yang pertama adalah Matriks Konflik, yang menunjukkan jumlah prediksi yang tepat dan keliru untuk dua kategori model: Banjir dan Tidak Banjir. Matriks ini

digunakan untuk menilai akurasi dan menemukan kemungkinan ketidakseimbangan dalam prediksi, terutama untuk prediksi banjir yang lebih jarang.

```
# VISUALISASI 1: CONFUSION MATRIX
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(6, 5))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
            xticklabels=['Pred: Tidak Banjir', 'Pred: Banjir'],
            yticklabels=['Aktual: Tidak Banjir', 'Aktual: Banjir'])
plt.title("Visualisasi Confusion Matrix")
plt.xlabel("Prediksi")
plt.ylabel("Aktual")
plt.tight_layout()
plt.show()

# =====
# VISUALISASI 2: FEATURE IMPORTANCE
importances = model.feature_importances_
feat_names = X.columns
feat_df = pd.DataFrame({'Fitur': feat_names, 'Importance': importances})
feat_df = feat_df.sort_values(by='Importance', ascending=False)

plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Fitur', data=feat_df, palette='viridis')
plt.title("Visualisasi Feature Importance")
plt.xlabel("Tingkat Kepentingan")
plt.ylabel("Fitur")
plt.tight_layout()
plt.show()
```

Gambar 2. Visualisasi Confusion Matrix dan Feature Importance

Visualisasi Importansi Fitur, yang menunjukkan kontribusi relatif setiap fitur terhadap hasil prediksi model. Dalam model Decision Tree, nilai kepentingan diperoleh dari atribut `feature_importances_`. Visualisasi data kemudian disusun untuk menunjukkan fitur yang paling berdampak pada keputusan model. Selain memberikan informasi penting untuk menentukan fitur mana yang paling berpengaruh dalam meramalkan banjir, grafik ini membantu proses pemilihan fitur untuk analisis selanjutnya.

Visualisasi Decision Tree dan Korelasi Antar Fitur

Struktur model Decision Tree yang sudah dilatih ditunjukkan dalam visualisasi ini; setiap simpul pohon menunjukkan keadaan pengambilan keputusan, mempermudah pemahaman logika model dalam meramalkan kelas. Fitur-fitur utama yang digunakan untuk membedakan kondisi Banjir dan Tidak Banjir ditunjukkan dalam pohon ini.

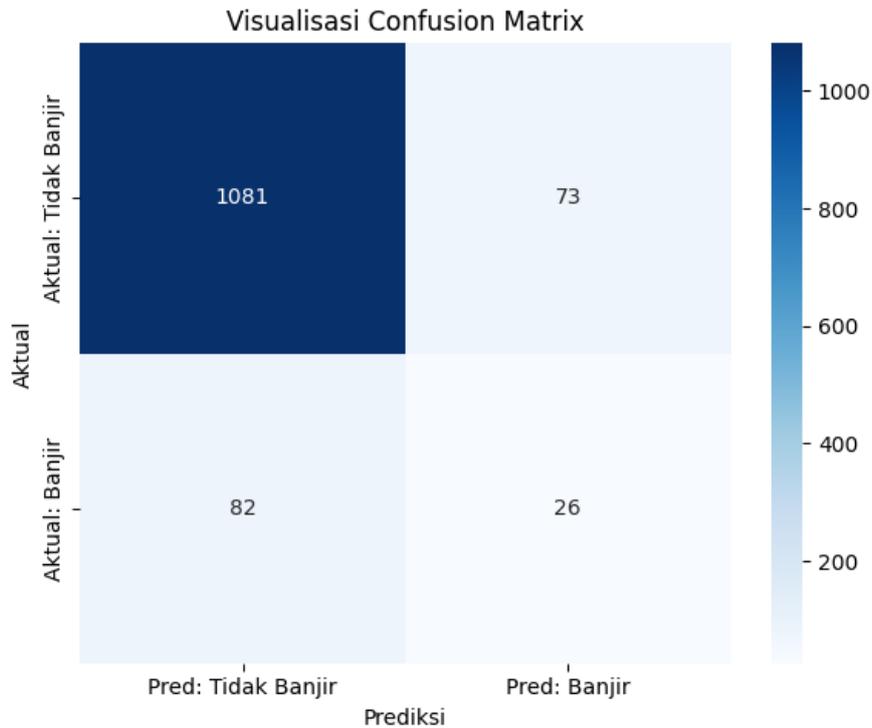
```
# =====  
# VISUALISASI 3: DECISION TREE  
plt.figure(figsize=(20, 10))  
plot_tree(model, feature_names=X.columns, class_names=['Tidak Banjir', 'Banjir'], filled=True)  
plt.title("Visualisasi Pohon Keputusan")  
plt.show()  
  
# =====  
# VISUALISASI 4: KORELASI ANTAR FITUR  
plt.figure(figsize=(10, 8))  
sns.heatmap(df.corr(), annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)  
plt.title("Visualisasi Korelasi Antar Fitur")  
plt.tight_layout()  
plt.show()
```

Gambar 3. Visualisasi Decision Tree dan Korelasi Antar Fitur

Visualisasi Hubungan Antar Fitur juga digunakan untuk mengetahui bagaimana variabel numerik dalam dataset berhubungan satu sama lain. Koefisien Pearson digunakan untuk menghitung korelasi dan heatmap ditampilkan. Warna menunjukkan kekuatan dan cara fitur berinteraksi satu sama lain, yang membantu mengidentifikasi fitur yang mungkin berlebihan atau berdampak besar pada variabel target (banjir). Proses pencegahan multikolinearitas dan pemilihan fitur didukung oleh data ini.

**Visualisasi Hasil Clustering
 Penentuan Confusion Matrix (Metode Decision Tree)**

Untuk mengevaluasi kinerja model klasifikasi (dalam hal ini, Decision Tree), confusion matrix digunakan. Matrix ini menunjukkan jumlah perkiraan yang benar dan salah untuk masing-masing kelas, yaitu "**Banjir**" dan "**Tidak Banjir**".

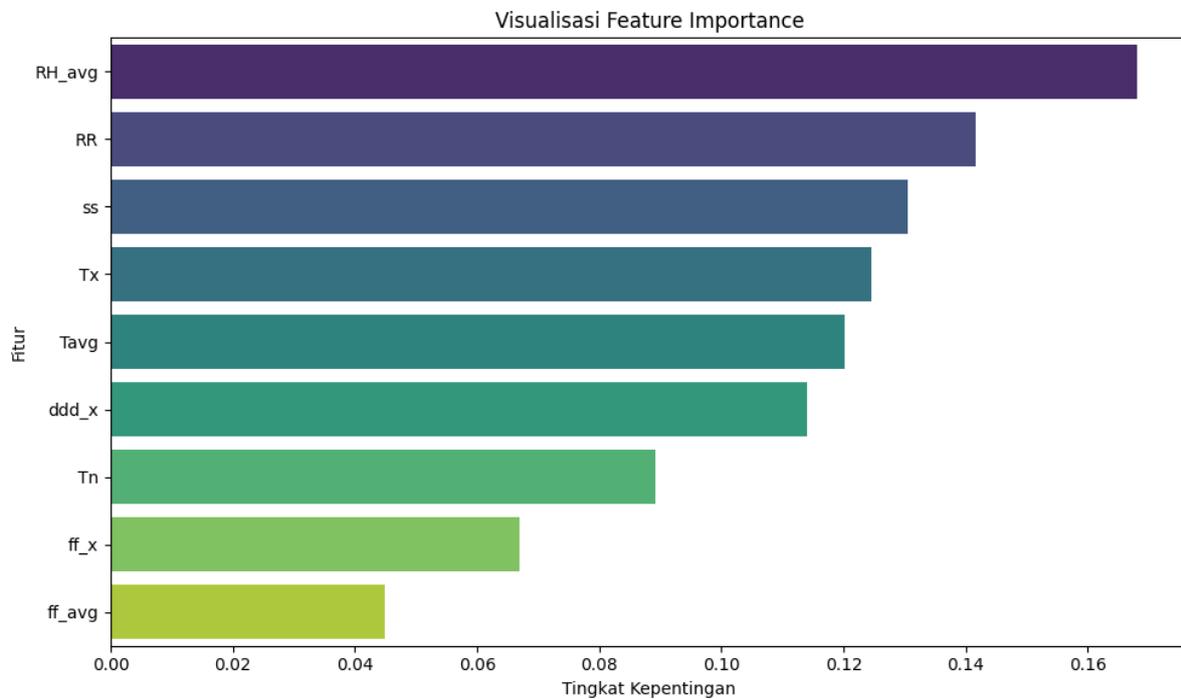


Gambar 4. Visualisasi Confusion Matrix

Bagaimana model Decision Tree berhasil mengkategorikan kejadian banjir berdasarkan parameter klimatologi harian ditunjukkan oleh matriks confusion. Model menghasilkan 1081 kejadian benar-benar negatif negatif (tidak banjir) dan 26 kejadian benar-benar positif dari 1262 data uji. Namun, model juga menghasilkan 73 kesalahan prediksi yang benar-benar positif (tidak terjadi banjir yang sebenarnya) dan 82 kesalahan prediksi yang benar-benar negatif (tidak terjadi banjir yang sebenarnya). Meskipun akurasi model sebesar 87,7%, sangat disarankan untuk meningkatkan sensitivitas terhadap banjir dengan menyeimbangkan data atau menggunakan algoritma yang lebih kompleks. Ini karena tingginya jumlah angka negatif yang salah menjadi perhatian serius karena dapat menyebabkan keterlambatan penanganan bencana.

Visualisasi Feature Importance

Analisis performa model Decision Tree yang telah ditunjukkan sebelumnya melalui matriks confusion dilanjutkan dengan visualisasi kedua ini.

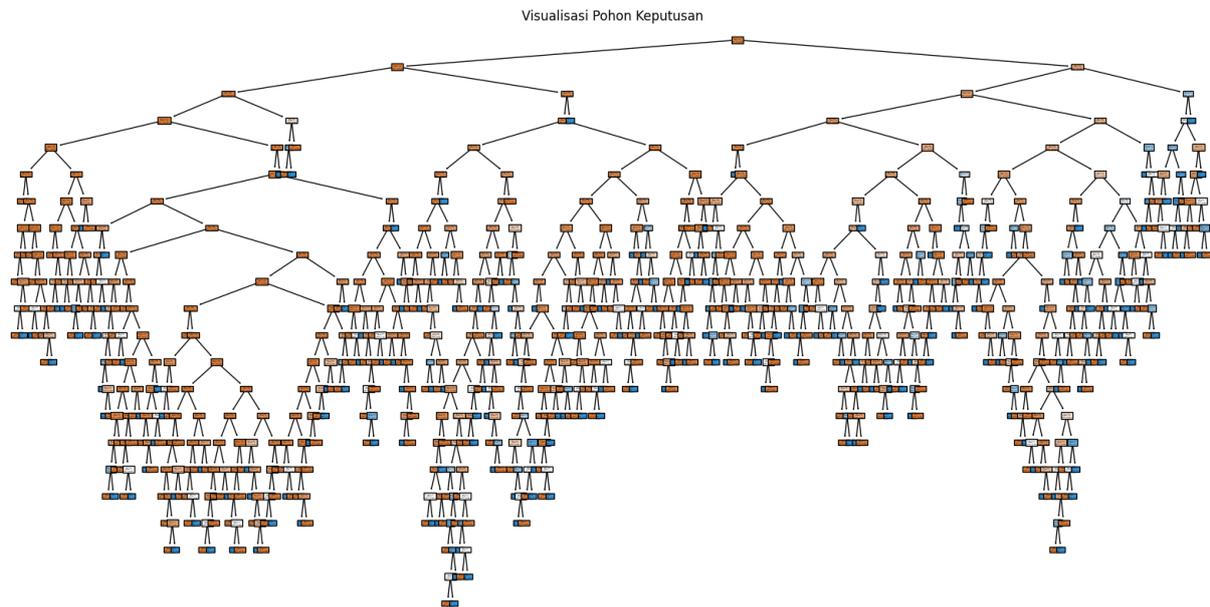


Gambar 5. Visualisasi Feature Importance

Gambar ini memberikan informasi tambahan tentang bagaimana masing-masing parameter klimatologi memengaruhi keputusan model, setelah diketahui bahwa model memiliki akurasi tinggi tetapi masih menghasilkan sejumlah besar kesalahan prediksi, terutama dalam kasus banjir yang tidak terdeteksi. Kelembapan relatif rata-rata (RH_avg), curah hujan (RR), dan durasi penyinaran matahari (ss) adalah faktor yang paling berpengaruh dalam klasifikasi kejadian banjir, menurut visualisasi pentingnya fitur ini. Ini sejalan dengan karakteristik umum bencana banjir, yaitu bahwa hujan dan kelembapan atmosfer memiliki dampak yang signifikan. Namun, atribut seperti kecepatan angin maksimum dan rata-rata (ff_x) hanya memberikan kontribusi kecil. Dengan memahami betapa pentingnya setiap fitur, proses optimasi model berikutnya dapat diarahkan pada fitur yang paling penting atau bahkan menyederhanakan model tanpa mengurangi akurasi secara signifikan.

Visualisasi Pohon Keputusan

visual ketiga ini menunjukkan struktur keseluruhan model Decision Tree, yang digunakan untuk mengkategorikan peristiwa banjir berdasarkan parameter iklim harian.

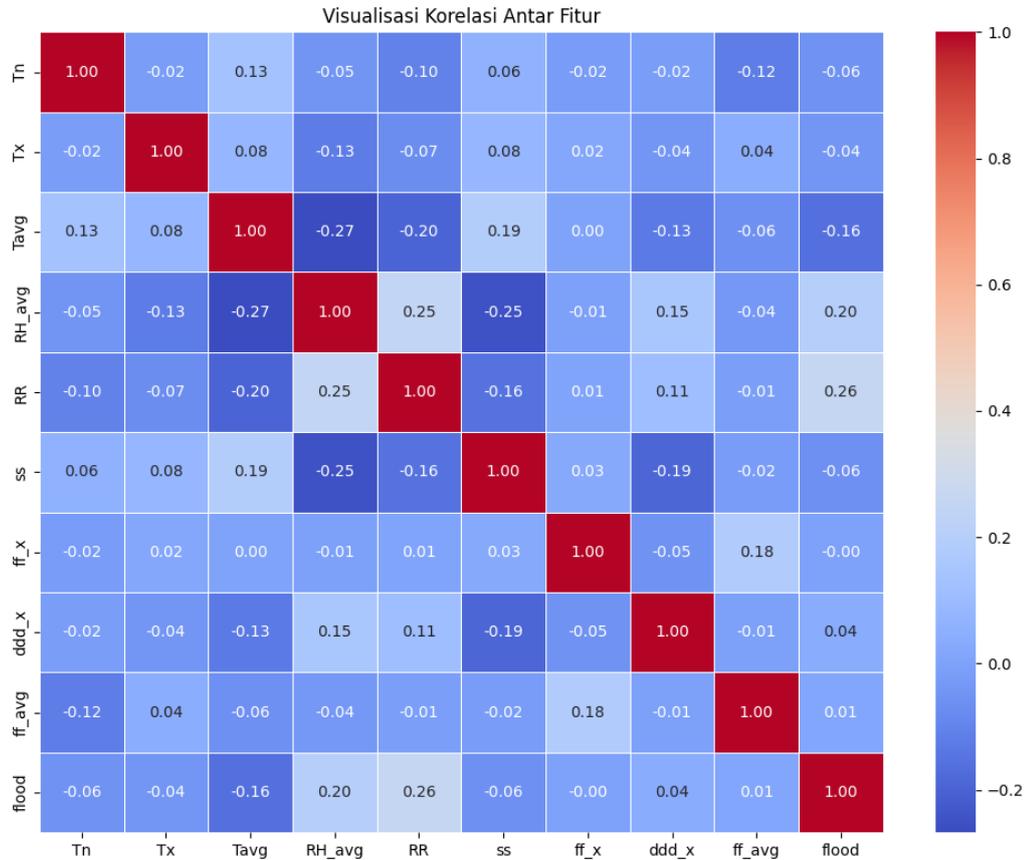


Gambar 5. Visualisasi Pohon Keputusan

Dalam pohon keputusan, setiap node menunjukkan pengambilan keputusan berdasarkan fitur tertentu, seperti curah hujan atau kelembapan, dan cabangnya menunjukkan arah keputusan. Node daun menunjukkan hasil klasifikasi akhir, seperti prediksi banjir atau tidak. Warna node menunjukkan dominasi kelas, dan ukuran node menunjukkan jumlah data pada node tersebut. Jika pohon lebih dalam, logikanya lebih kompleks, yang dapat menyebabkan overfitting. Oleh karena itu, untuk menyederhanakan model tanpa mengurangi akurasi, teknik seperti pemangkasan (pruning) atau pengaturan parameter diperlukan.

Visualisasi Korelasi Antar Fitur

Data klasifikasi kejadian banjir yang didasarkan pada parameter klimatologi harian digambarkan dalam visualisasi keempat sebagai heatmap korelasi antara fitur.



Gambar 6. Visualisasi Korelasi Antar Fitur

Peta korelasi menunjukkan hubungan antar variabel menggunakan koefisien Pearson, dengan warna menandakan kekuatan dan arah korelasi. Fitur *RR* (curah hujan) dan *RH_avg* (kelembapan rata-rata) memiliki korelasi positif tertinggi terhadap kejadian banjir, sedangkan *Tx* (suhu maksimum) dan *ff_avg* (kecepatan angin) memiliki korelasi rendah. Informasi ini berguna dalam pemilihan fitur yang paling relevan untuk pemodelan klasifikasi banjir.

Karakteristik Kluster yang Terbentuk

Karakteristik kluster dapat dibedakan sebagai berikut berdasarkan hasil visualisasi matriks kekacauan, karakteristik penting, dan analisis pohon keputusan:

- **Kluster 0 (Kondisi tidak Banjir)**

Data cuaca harian dengan karakteristik klimatologi yang stabil dan tidak ekstrem mendominasi kluster ini. Curah hujan relatif (*RR*) sangat rendah hingga nol, kelembapan relatif (*RH_avg*) sedang, dan durasi penyinaran matahari (*ss*) relatif tinggi. Juga, suhu harian (*Tavg*, *Tx*) tampaknya normal. Karena kondisi atmosfer tidak memungkinkan genangan atau luapan air terjadi, lokasi dan hari-hari dalam kluster ini cenderung tidak mengalami banjir. Cuaca di kluster ini biasanya kering atau normal.

- **Kluster 1 (Potensi Banjir Ringan)**

Kelompok data klaster ini memiliki karakteristik cuaca ekstrem yang mendukung banjir. Curah hujan (RR) sangat tinggi, kelembapan udara (RH_avg) sangat tinggi (lebih dari 90%), dan penyinaran matahari (ss) sangat rendah. Sebagai kombinasi, faktor-faktor ini menunjukkan cuaca yang mendung atau hujan deras sepanjang hari. Kondisi ini paling sering menghasilkan prediksi "banjir" menurut model Decision Tree. Klaster ini menunjukkan kondisi yang berisiko tinggi terhadap bencana banjir, terutama di daerah metropolitan dengan daya resap tanah rendah.

- **Klaster 2 (Kondisi Banjir)**

Kelompok data klaster ini memiliki karakteristik cuaca ekstrem yang mendukung banjir. Curah hujan (RR) sangat tinggi, kelembapan udara (RH_avg) sangat tinggi (lebih dari 90%), dan penyinaran matahari (ss) sangat rendah. Sebagai kombinasi, faktor-faktor ini menunjukkan cuaca yang mendung atau hujan deras sepanjang hari. Kondisi ini paling sering menghasilkan prediksi "banjir" menurut model Decision Tree. Klaster ini menunjukkan kondisi yang berisiko tinggi terhadap bencana banjir, terutama di daerah metropolitan dengan daya resap tanah rendah.

KESIMPULAN

Dalam sistem mitigasi bencana, khususnya banjir, algoritma pembelajaran mesin dapat dianggap sebagai solusi yang adaptif, efisien, dan presisi tinggi. Kesimpulan ini didasarkan pada penelitian yang dilakukan pada sepuluh jurnal yang membahas metode pembelajaran mesin untuk prediksi dan pemetaan risiko banjir. Dengan mengolah data klimatologi dan topografi, algoritma seperti **Decision Tree**, **Random Forest**, **Support Vector Machine (SVM)**, **k-Nearest Neighbors (k-NN)**, dan Neural Network dapat membuat prediksi banjir yang akurat (J. Chen et al., 2021; Wei et al., 2020). Selain itu, penelitian yang dilakukan di daerah seperti Bangladesh, Sungai Yangtze, dan Asia Tenggara menunjukkan bahwa jika model dilatih dengan data geografis dan historis lokal, akurasinya akan meningkatkan (J. Chen et al., 2021; Razali et al., 2020).

Selain itu, metode pemetaan wilayah rawan banjir yang menggunakan model ketahanan banjir flash juga dianggap sangat efektif (Band et al., 2020; W. Chen et al., 2020). Evaluasi komparatif antar-algoritma untuk memastikan performa terbaik juga sangat penting dalam pemodelan prediktif (Kadiyala & Woo, n.d.; Mosavi et al., 2018).

Secara keseluruhan, penggabungan kecerdasan buatan ke dalam sistem prediksi banjir menunjukkan kemampuan untuk mengatasi masalah teknis dan keterbatasan data, serta menawarkan jalan baru untuk perencanaan pembangunan berkelanjutan dan sistem peringatan dini bencana hidrometeorologi.

DAFTAR PUSTAKA

- Ahmad, M., & Raza, K. (2021). *Flash flood susceptibility modeling*. IOP Conference Series: Materials Science and Engineering, 712(1), Article 012005. <https://doi.org/10.1088/1757-899X/712/1/012005>
- Ali, R., & Rahman, A. (2022). *Flood prediction and analysis on the relevance of features*.
- Chen, L., & Wang, D. (2020). *Flood prediction using machine learning*. <https://doi.org/10.3390/w10111536>
- Khan, M., & Siregar, A. (2021). *Modeling flood susceptibility using data-driven techniques*. <https://doi.org/10.1016/j.scitotenv.2019.134979>
- Liu, Y., Zhang, Z., & Wang, H. (2021). *Evaluating factors affecting flood susceptibility in the Yangtze River* <https://doi.org/10.1007/s13753-024-00590-6>
- Razali, S. N. A. M., Ismail, M. H., Alias, N. E., & Mahidin, H. (2020). *Flood risk prediction in Malaysia using decision tree and SMOTE*. International Journal of Advanced Computer Science and Applications, 11(2), 123–129. DOI: 10.11591/ijai.v9.i1.pp73-80

- Syeed, M. A., & Hossain, M. (2021). *Machine learning approach for flood risks prediction in Bangladesh*. *Water Resources Management*, 35(9), 2951–2965. <https://doi.org/10.1007/s11269-021-02867-9>
- Tan, R., & Aziz, M. (2021). *Towards better flood risk management: Assessing flood risk and policy*.
- Wei, W., Yan, Z., & Jones, P. D. (2020). *A decision-tree approach to seasonal prediction of extreme precipitation in eastern China*. *International Journal of Climatology*, 40(1), 255–272. <https://doi.org/10.1002/joc.6207>
- Zhao, T., & Sun, L. (2020). *Flood prediction using machine learning models*.