

Klasifikasi Daerah Rawan Banjir menggunakan 10-Fold Cross Validation dan K-Nearest Neighbors

Classification of Flood-Prone Areas Using 10-Fold Cross Validation and K-Nearest Neighbors

¹Adyatma Andhika Bagaskara*, ²Kristoko Dwi Hartomo

Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana
Jalan Dr. O. Notohamidjojo No.1 - 10, Blotongan, Kec. Sidorejo, Kota Salatiga, Jawa Tengah 50715,
Indonesia

*e-mail: 682019079@student.uksw.edu

(*received:* 31 Oktober 2023, *revised:* 5 November 2023, *accepted:* 6 November 2023)

Abstrak

Banjir terjadi ketika volume air melebihi kapasitas penampungan atau aliran normal saluran air, seperti sungai, danau, kanal, atau sistem drainase. Pada tahun 2021 dan 2022, terjadi 88 bencana banjir di Semarang. Faktor-faktor yang memengaruhi terjadinya bencana banjir termasuk curah hujan, kemiringan lahan, dan indeks tutupan vegetasi (NDVI). Tujuan utama penelitian ini adalah untuk mengidentifikasi dan mengklasifikasi daerah-daerah di Semarang yang rentan terhadap banjir dengan memetakannya berdasarkan curah hujan, indeks vegetasi, dan kemiringan lahan. Status yang diklasifikasikan adalah daerah rawan banjir dan daerah tidak rawan banjir. Evaluasi spasial daerah rawan banjir berdasarkan GIS dilakukan dengan menggunakan algoritma klasifikasi K-Nearest Neighbor (K-NN) dengan bahasa pemrograman R untuk mencapai tujuan ini. Pengujian k optimal menggunakan metode 10-Fold Cross Validation mengungkapkan bahwa akurasi tertinggi sistem tercapai dengan metode K-NN pada k=7, mencapai 86%. Temuan dari penelitian ini menunjukkan bahwa delapan kecamatan di Semarang rentan terhadap banjir, sedangkan 167 daerah di Semarang tidak menunjukkan kerentanan terhadap banjir. Potensi daerah rawan banjir direpresentasikan dengan pembuatan peta kerawanan banjir Semarang menggunakan perangkat lunak Quantum GIS.

Kata kunci: Banjir, K-NN, QGIS, 10-fold cross-validation, Pemetaan

Abstract

Floods occur when the volume of water exceeds the capacity of containment or the normal flow of water channels, such as rivers, lakes, canals, or drainage systems. In 2021 and 2022, there were 88 flood disasters in Semarang. Factors influencing the occurrence of flood disasters include rainfall, land slope, and the vegetation cover index (NDVI). The main objective of this research is to identify and classify areas in Semarang that are susceptible to flooding by mapping them based on rainfall, vegetation index, and land slope. The classified statuses are flood-prone areas and non-flood-prone areas. A spatial assessment of areas prone to flooding based on GIS is performed using the K-Nearest Neighbor (K-NN) classification algorithm with the R programming language to accomplish this objective. Optimal k testing using the 10-fold cross-validation method reveals that the highest accuracy of the system is achieved with the K-NN method at k=7, reaching 86%. The findings from this study suggest that eight sub-district areas within Semarang are susceptible to flooding, whereas 167 areas in Semarang do not exhibit susceptibility to floods. The potential areas prone to flooding are represented by making a Semarang flood susceptibility map using Quantum GIS software.

Keywords: Flood, K-NN, QGIS, 10-fold cross-validation, Mapping

1 PENDAHULUAN

Semarang adalah salah satu kota di Indonesia yang sering terkena bencana banjir. Berdasarkan data Indeks Risiko Bencana Indonesia (IRBI) [1], Semarang memiliki tingkat risiko tinggi terhadap bencana banjir. Pada tahun 2021, Semarang mengalami 432 bencana, di antaranya 88 adalah banjir,

<http://sistemasi.ftik.unisi.ac.id>

menempati peringkat kedua setelah tanah longsor dalam tujuh jenis bencana yang terjadi. Sebagai salah satu kota yang rawan banjir di Indonesia, pemetaan akurat diperlukan untuk mengidentifikasi daerah yang rentan terhadap banjir.

Salah satu langkah untuk mengurangi dan mencegah kerugian akibat banjir adalah dengan melakukan evaluasi risiko banjir. Evaluasi risiko banjir adalah alat yang digunakan untuk menilai sejauh mana kemungkinan dan besarnya dampak yang dapat ditimbulkan oleh banjir di suatu wilayah. Evaluasi risiko banjir merupakan dasar yang diperlukan untuk memastikan koordinasi dan efektivitas upaya penanggulangan bencana [2].

Dalam penelitian ini, Sistem Informasi Geografis (SIG) menjadi alternatif untuk menggambarkan hasil penilaian risiko banjir, terutama di Kota Semarang melalui representasi peta spasial. SIG adalah suatu sistem komputer yang digunakan untuk mengelola data lingkungan dengan fokus pada unsur geografis. Prinsip-prinsip yang ada dalam SIG dapat mendukung dalam pembuatan peta dan analisis risiko banjir di wilayah yang sedang diobservasi. [2].

Algoritma K-Nearest Neighbors (K-NN) dan perangkat lunak *Quantum Geographic Information System* (QGIS) digunakan sebagai alat analisis. Algoritma K-NN adalah salah satu algoritma pembelajaran mesin yang digunakan untuk klasifikasi. K-Nearest Neighbor membuat prediksi melalui *supervised learning*, di mana hasil dari suatu contoh kueri ditentukan berdasarkan mayoritas kategori tetangga terdekatnya[3]. Algoritma ini bekerja dengan mencari tetangga terdekat dari titik data yang diberikan. Dalam konteks pemetaan daerah rawan banjir, algoritma K-NN dapat memetakan daerah yang rentan terhadap banjir berdasarkan data yang terkumpul.

Tujuan dari penelitian ini yaitu menggunakan data yang digunakan meliputi data curah hujan, data kemiringan lahan, dan data NDVI untuk menentukan daerah rawan banjir di Kota Semarang. Data ini akan diproses menggunakan algoritma K-NN dan QGIS untuk memetakan daerah yang rawan. Keunikan penelitian ini terletak pada optimasi K-NN dengan metode *10-Fold Cross-Validation* untuk meningkatkan akurasi hasil klasifikasi.

2 TINJAUAN LITERATUR

Untuk menunjang penelitian ini, perlu merujuk kepada literatur-literatur yang sudah ada sebelumnya. Literatur tersebut mencakup hal-hal yang dapat digunakan dalam penelitian ini, seperti K-NN, *10-Fold Cross-Validation*, kemiringan lahan, NDVI, dan curah hujan.

2.1 K-Nearest Neighbors (K-NN)

K-Nearest Neighbor (K-NN) adalah sebuah algoritma machine learning yang digunakan untuk tugas klasifikasi dan regresi [3]. Algoritma ini didasarkan pada konsep bahwa objek-objek dengan fitur yang serupa cenderung mengelompok dalam ruang fitur. K-Nearest Neighbor (K-NN) termasuk dalam kategori pembelajaran berbasis contoh. Algoritma ini juga dianggap sebagai teknik pembelajaran yang malas (*lazy learning*). K-NN melibatkan pencarian sekelompok objek k dalam data pelatihan yang paling dekat dengan objek dalam data uji atau data latih [4].

K-NN telah diimplementasikan oleh Tempola [5] untuk menentukan status gunung berapi dengan menggunakan validasi silang K-Fold untuk pengujian akurasi algoritma. Berdasarkan pengujian yang dilakukan pada sistem tersebut, akurasi rata-rata sebesar 63,68% diperoleh dengan deviasi standar sebesar 7,47%. K-NN juga telah diimplementasikan oleh Pasaribu [6] untuk pemantauan bencana alam menggunakan data media sosial. Para peneliti menyimpulkan bahwa nilai pengujian matriks kebingungan terbaik dengan metrik Jaccard adalah 86%. Selain itu, algoritma K-NN juga diimplementasikan dalam studi oleh Cumel [7] untuk memprediksi banjir di Jakarta. Data yang diprediksi dengan akurat terbukti dalam *confussion matrix*, dengan nilai akurasi tertinggi sebesar 88,9% dengan K=5. Selain untuk prediksi banjir, K-NN juga telah digunakan dalam penelitian oleh Fitri untuk mengklasifikasikan kelainan sel darah putih dengan akurasi terbaik mencapai 94,3% dengan k = 23 [8].

2.2 10-Fold Cross-Validation

Validasi silang (*cross-validation*) adalah metode validasi model yang digunakan untuk menilai sejauh mana hasil analisis statistik berlaku untuk dataset yang secara independen terpisah. Teknik ini

utamanya digunakan untuk prediksi dengan model-model dan untuk mengevaluasi tingkat akurasi model-model prediksi saat diterapkan dalam situasi dunia nyata [9].

Menggunakan validasi silang K-Fold memberikan keuntungan dalam mengidentifikasi model dengan akurasi tertinggi. Hal ini terjadi karena data dibagi secara acak, memungkinkan kita untuk mengamati komposisi model yang paling optimal [10].

2.3 Kemiringan Lahan

Kemiringan lereng memiliki dampak pada jumlah dan kecepatan air permukaan yang mengalir, drainase permukaan, penggunaan lahan, dan erosi. Jika lerengnya datar, aliran air permukaan akan lambat, meningkatkan risiko genangan atau banjir. Sebaliknya, lereng yang curam akan mempercepat aliran air permukaan, mengalirkan hujan dengan cepat dan mengurangi risiko banjir [11].

Peta kemiringan lahan didapatkan dari proses pengolahan data *Digital Elevation Model* (DEM) suatu area dengan analisa *slope* menggunakan *software* QGIS. Data DEM yang sudah diolah kemudian dikelompokkan ke kelas-kelas untuk digunakan sebagai parameter analisa daerah rawan banjir.

2.4 Normalized Difference Vegetation Index (NDVI)

NDVI merupakan indeks vegetasi yang menandakan tingkat kehijauan suatu wilayah yang diolah dari pengolahan beberapa kanal sensor satelit.

NDVI dihitung dengan menggunakan rumus $NDVI = (NIR - RED) / (NIR + RED)$, di mana NIR (Near-Infrared) dan RED (Merah) mewakili jumlah reflektansi dalam kanal inframerah-dekat dan kanal merah yang dipantulkan oleh vegetasi dan diukur oleh sensor satelit. Formula tersebut didasarkan pada fakta bahwa klorofil menyerap cahaya merah sedangkan mesofil memantulkan cahaya inframerah dekat. Nilai NDVI berkisar dari -1 ke +1, di mana nilai negatif menandakan tidak adanya vegetasi [12].



Gambar 1. Rentang Nilai NDVI

Gambar 1 menjelaskan bahwa daerah yang memiliki nilai NDVI mendekati 1 menandakan wilayah dengan vegetasi tinggi (hutan lebat), sedangkan wilayah yang memiliki nilai NDVI mendekati 0 maka menandakan wilayah dengan vegetasi rendah atau tidak ada vegetasi (perairan atau tanah bebatuan) [13].

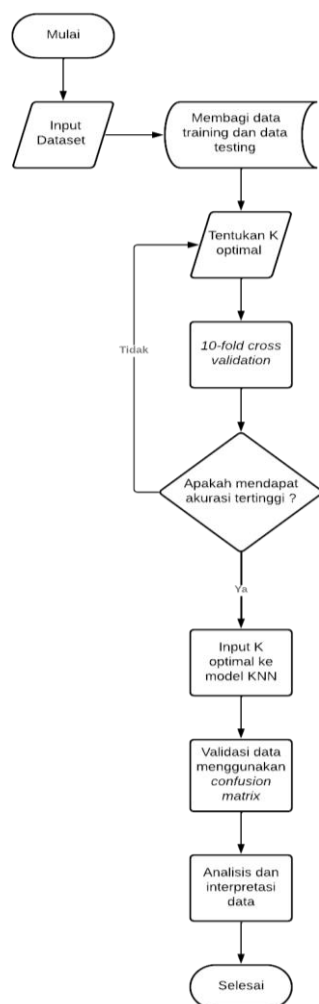
2.5 Curah Hujan

Curah hujan merupakan besarnya volume hujan yang jatuh pada satu area pada periode tertentu.. Curah hujan sangat mempengaruhi terjadinya potensi bencana banjir, jika suatu daerah memiliki curah hujan tinggi maka daerah tersebut berpotensi tinggi menghadapi bencana banjir sedangkan jika daerah tersebut memiliki curah hujan yang rendah maka daerah tersebut berpotensi rendah dalam menghadapi bencana banjir [11]. Satuan curah hujan dinyatakan dalam millimeter dan dihitung dengan membagi volume air hujan dengan luas alas, yang artinya jika suatu lokasi memiliki curah hujan sebesar 10mm maka lokasi tersebut digenangi air setinggi 10 mm.

Metode IDW merupakan salah satu metode interpolasi yang dapat digunakan untuk membuat peta curah hujan yang mudah dipahami dan mudah diimplementasikan. Metode IDW digunakan untuk mengetahui nilai dari suatu titik sampel yang belum diketahui nilainya [14]. Keunggulan utama metode interpolasi IDW adalah kemampuannya untuk mengatur ciri-ciri interpolasi dengan membatasi penggunaan titik-titik masukan dalam proses tersebut. Titik-titik yang berjauhan dari titik sampel dan dianggap memiliki hubungan spasial yang rendah atau bahkan tidak ada hubungan spasial bisa dikecualikan dari perhitungan [15].

3 METODE PENELITIAN

Penelitian ini menggunakan algoritma K-Nearest Neighbor dengan 10-Fold Cross Validation untuk menemukan nilai K optimal. Diagram alur untuk metode K-NN digambarkan dalam Gambar 2.



Gambar 2. Flowchart Pengujian K-NN

Gambar 2 menjelaskan klasifikasi daerah rawan banjir di Kota Semarang dimulai dengan langkah pertama yaitu memasukkan dataset ke dalam sistem. Selanjutnya, dataset ini dibagi menjadi dua bagian, yakni data uji dan data latih. Kemudian, untuk menemukan nilai k yang optimal, dilakukan pencarian menggunakan metode *10-fold cross validation*. Setelah mendapatkan nilai k optimal dengan akurasi tertinggi, langkah selanjutnya adalah menggunakan model K-Nearest Neighbors (KNN) dengan nilai k tersebut untuk melakukan pengujian. Untuk menguji validitas data dan menganalisis hasilnya, digunakan *confusion matrix*. Hasil dari model KNN ini kemudian dianalisis dan diinterpretasikan untuk menciptakan peta yang menunjukkan daerah-daerah rawan banjir di kota Semarang.

3.1 Jenis dan Sumber Data

Penelitian ini menggunakan data *band 5* dan *band 4* citra Landsat 8 OLI dengan jalur/baris 120/065 untuk diolah menjadi NDVI diperoleh dari Januari 2021 hingga Desember 2022, yang diambil dari United States Geological Survey (USGS). Data curah hujan harian diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika Indonesia (BMKG), di mana data curah hujan harian diambil mulai dari Januari 2021 hingga Desember 2022. Data kemiringan lahan *Digital Elevation Model*

(DEM) dapat digunakan untuk membuat peta kemiringan lahan. Peta DEM diunduh dari situs web Geospasial Indonesia.

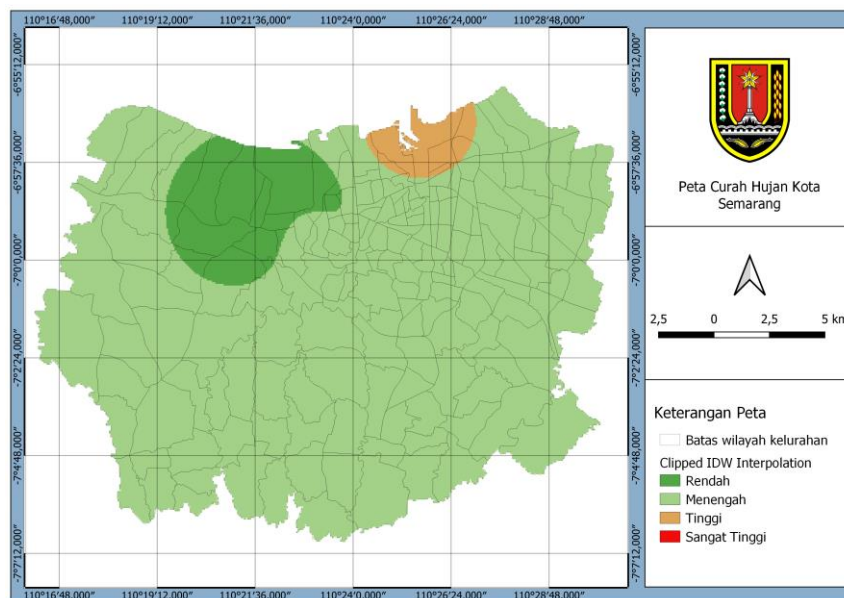
3.2 Pengolahan Data

Data NDVI, curah hujan dan kemiringan lahan yang sudah diolah dari citra satelit dikelompokkan menjadi 175 dataset sesuai dengan jumlah kelurahan yang ada di Kota Semarang kemudian dibagi menjadi 124 data latih dan 51 data uji untuk dimasukkan ke dalam algoritma K-NN. *10-Fold Cross Validation* dilakukan sebelum memulai algoritma untuk mencari K dengan akurasi tertinggi. Hasil dari pengolahan data menggunakan K-NN diperiksa ulang dengan *confusion matrix* untuk mengetahui tingkat akurasi algoritma.

4 HASIL DAN PEMBAHASAN

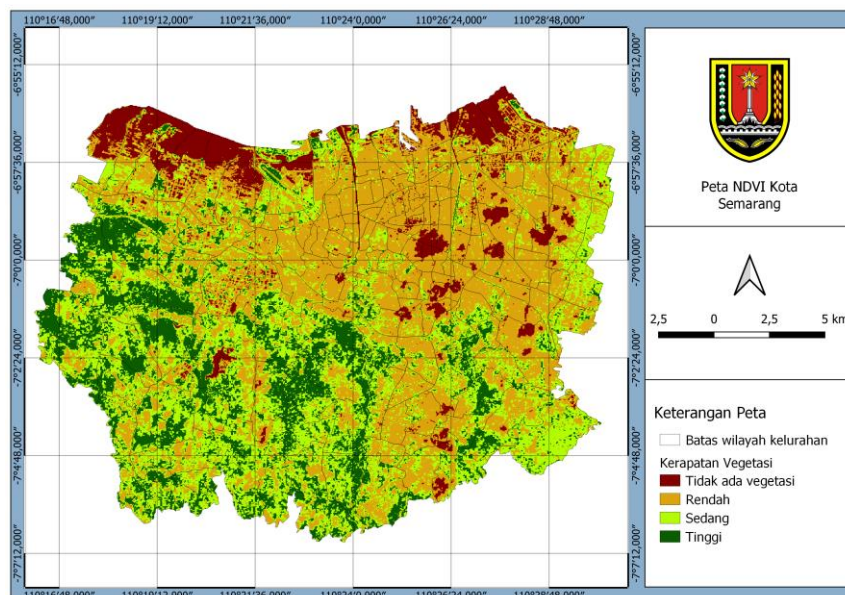
Pengujian dengan algoritma K-Nearest Neighbor (K-NN) menggunakan bahasa pemrograman R dilakukan dengan memulai inialisasi berbagai nilai k, kemudian memilih nilai k yang memberikan akurasi tertinggi untuk proses validasi data menggunakan 10 fold-cross validation. Nilai K dengan akurasi tertinggi dipilih untuk menjalankan model K-NN dan diuji menggunakan confusion matrix untuk mengetahui keakuratan model K-NN.

Pengujian sistem menggunakan metode klasifikasi K-NN melibatkan 3 data, yaitu data curah hujan, data kemiringan lahan dan data NDVI. Sementara itu hasil pengolahan diklasifikasikan menjadi dua kategori yaitu, wilayah terprediksi banjir dan wilayah terprediksi tidak banjir.



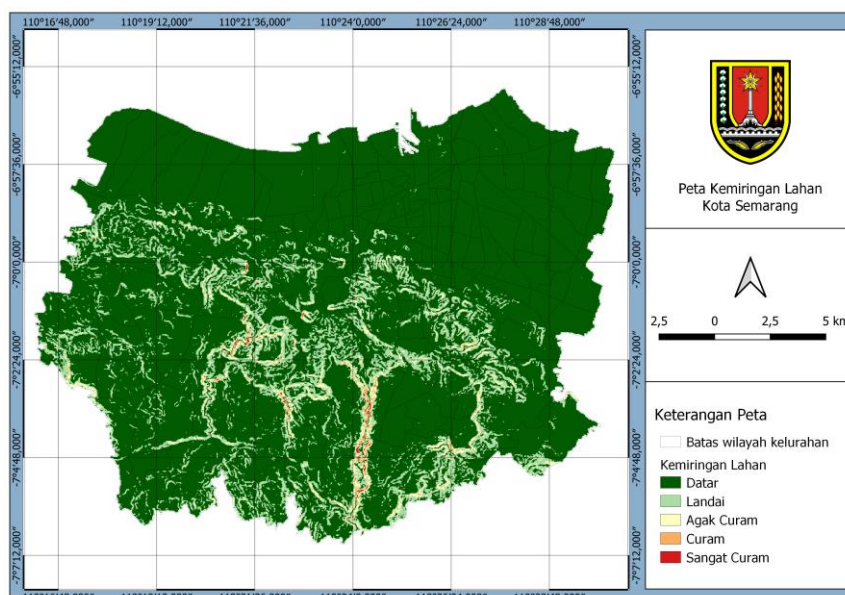
Gambar 3. Peta Curah Hujan Kota Semarang

Rata-rata curah hujan di Semarang diperoleh dari pengolahan data curah hujan oleh BMKG menggunakan metode IDW. Dapat dilihat pada gambar 3, curah hujan di Semarang dibagi menjadi empat klasifikasi: daerah dengan curah hujan rendah, daerah dengan curah hujan menengah, daerah dengan curah hujan tinggi, dan daerah dengan curah hujan sangat tinggi



Gambar 4. Peta NDVI Kota Semarang

Gambar 4 menunjukkan indeks vegetasi di Semarang, di mana sebagian besar wilayah memiliki vegetasi dengan kepadatan rendah. Wilayah-wilayah dengan vegetasi berkepadatan rendah terutama meliputi bagian-bagian bawah Semarang, seperti kecamatan Tugu, Semarang Barat, Semarang Utara, Semarang Tengah, Semarang Timur, Gayamsari, Pedurungan, Genuk, Gajahmungkur, dan Candisari.



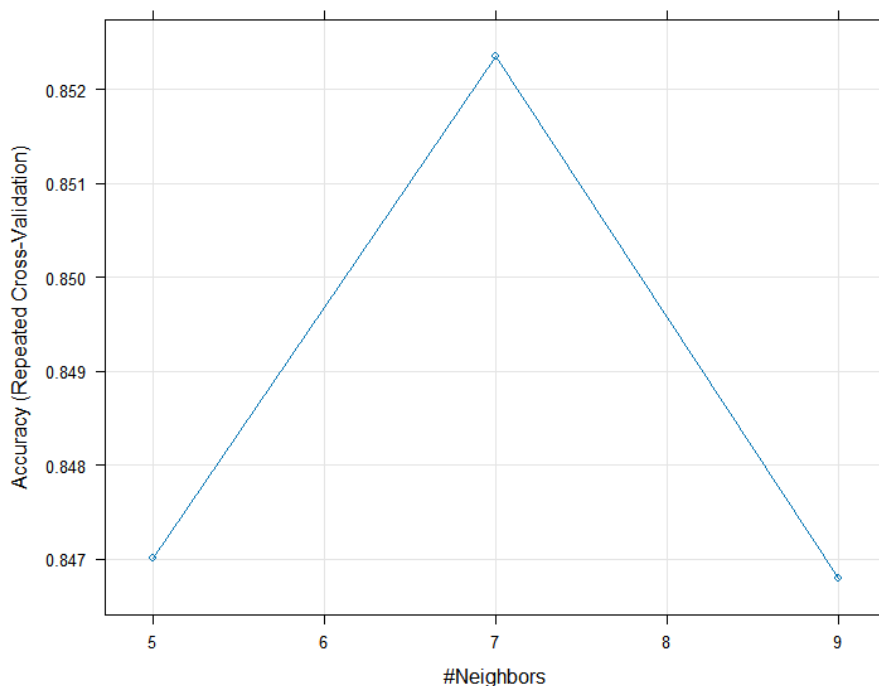
Gambar 5. Peta Kemiringan Lahan Kota Semarang

Gambar 5 mengindikasikan bahwa dari 175 lingkungan di wilayah Semarang, sebagian besar memiliki lereng yang datar, dengan hanya sebagian kecil memiliki lereng yang lebih curam.

Dalam penelitian ini, sebanyak 175 dataset per lingkungan dikumpulkan. Sebelum pengujian dengan validasi data, pengujian dilakukan dengan menerapkan berbagai nilai parameter 'k' dan hasil akurasi untuk setiap parameter 'k' diuji. Parameter 'k' dengan akurasi tertinggi dipilih untuk pengujian lanjutan menggunakan data validasi.

Data dibagi menjadi 124 data latih dan 51 data uji untuk pengujian dengan berbagai nilai 'k'. Metode *10-fold cross validation* digunakan untuk menguji variasi 'k'. Hasil akurasi untuk pengujian setiap 'k' adalah sebagai berikut: k=5 dengan akurasi 84%, k=7 dengan akurasi 85%, dan k=9 dengan

akurasi 84%. Nilai 'k' dengan akurasi tertinggi digunakan untuk menjalankan model K-NN, maka nilai 'k' yang digunakan untuk model adalah k=7.



Gambar 6. Grafik Akurasi Parameter K

Pada gambar 6, hasil akurasi parameter 'k' menunjukkan nilai tertinggi pada angka 7, yaitu sebesar 85%. Oleh karena itu, k=7 digunakan untuk model K-NN dalam prediksi bencana banjir. Data NDVI, curah hujan, dan kemiringan lahan dari 175 lingkungan diproses menggunakan model K-NN dengan k=7. Analisis *confusion matrix* digunakan untuk mengukur kinerja model.

```
Confusion Matrix and Statistics

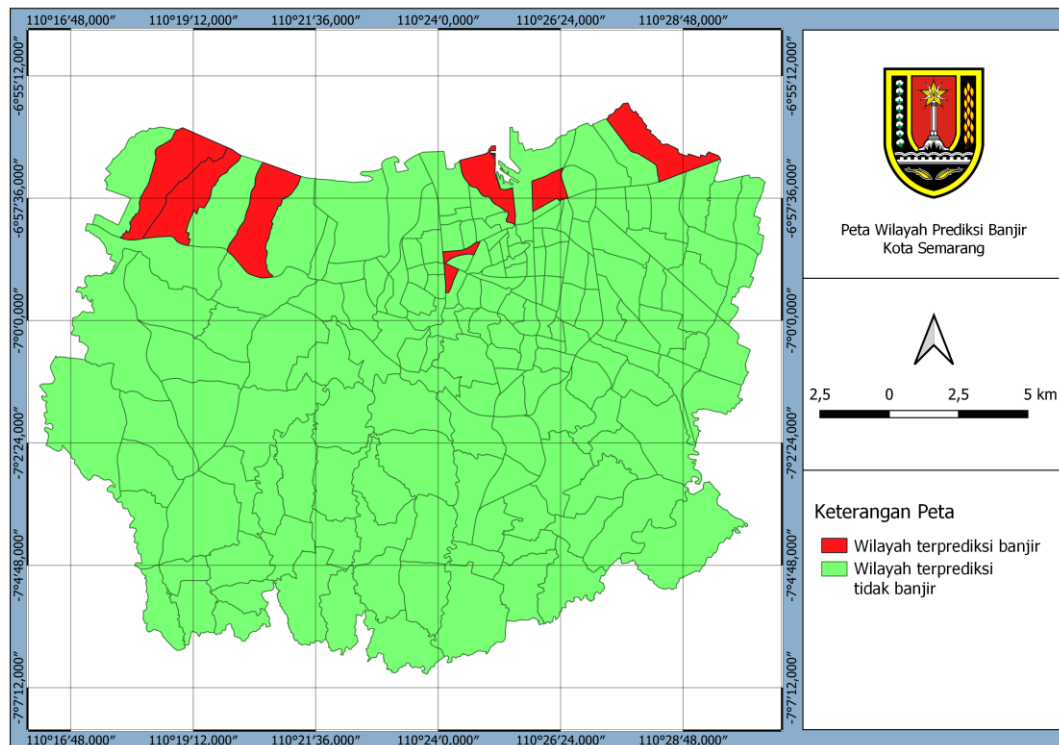
      Reference
Prediction FLOOD No Flood
FLOOD      5      3
No Flood   21     146

      Accuracy : 0.8629
      95% CI : (0.8028, 0.9101)
No Information Rate : 0.8514
P-Value [Acc > NIR] : 0.3836104
```

Gambar 7. Confusion Matrix

Data pada Gambar 7 menunjukkan tingkat akurasi model sebesar 86%, dengan delapan area yang diidentifikasi sebagai wilayah terprediksi banjir dan 167 sebagai wilayah yang terprediksi tidak banjir.

Hasil prediksi banjir untuk setiap kelurahan di Kota Semarang direpresentasikan secara visual dengan membuat peta menggunakan perangkat lunak *open source Quantum GIS*. Visualisasi peta prediksi banjir dapat dilihat pada Gambar 8.



Gambar 8. Peta Prediksi Banjir Kota Semarang

Gambar 8 mengilustrasikan bahwa wilayah yang berwarna merah merupakan wilayah terprediksi banjir, sementara wilayah yang berwarna hijau merupakan wilayah terprediksi tidak banjir. Wilayah yang terprediksi banjir meliputi Kecamatan Bandarharjo, Kecamatan Bulustalan, Kecamatan Karang Anyar, Kecamatan Kemijen, Kecamatan Mangkang Wetan, Kecamatan Mangunharjo, Kecamatan Pindrikan Lor, dan Kecamatan Trimulyo sementara wilayah yang terprediksi tidak banjir mencakup wilayah di luar wilayah terprediksi banjir.

5 KESIMPULAN

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa penelitian ini berhasil memprediksi daerah rawan banjir di Kota Semarang. Algoritma *K-Nearest Neighbor* (K-NN) dapat menentukan klasifikasi banjir dengan mempertimbangkan hubungan antara curah hujan, NDVI, dan kemiringan lahan. Nilai optimal 'k' dalam algoritma K-NN ditentukan berdasarkan akurasi rata-rata melalui *10-fold cross-validation*, yang menghasilkan akurasi tertinggi pada $k=7$, mencapai 85%. Hasil pengolahan data curah hujan, NDVI, dan kemiringan lahan untuk setiap kecamatan di Semarang menggunakan algoritma K-NN menunjukkan bahwa daerah rawan banjir meliputi Kecamatan Bandarharjo, Kecamatan Bulustalan, Kecamatan Karang Anyar, Kecamatan Kemijen, Kecamatan Mangkang Wetan, Kecamatan Mangunharjo, Kecamatan Pindrikan Lor, dan Kecamatan Trimulyo dengan akurasi data sebesar 86% sebagaimana diukur dengan analisa *confusion matrix*. Untuk penelitian selanjutnya, disarankan untuk menggunakan data yang lebih baru dan menggabungkan parameter tambahan seperti aliran sungai, NDWI, dan SAVI untuk meningkatkan akurasi model K-NN dalam memprediksi daerah rawan banjir. Untuk membuat peta potensi banjir yang lebih rinci dan spesifik, pemetaan daerah rawan banjir juga dapat dilakukan dalam skala yang lebih kecil, seperti pada tingkat RT dan RW.

REFERENSI

- [1] A. W. Adi et al., "Indeks risiko bencana Indonesia tahun 2021," pp. 11–3, 2022. \
- [2] A. T. Ujung, A. L. Nugraha, and H. S. Firdaus, "Kajian Pemetaan Risiko Bencana Banjir Kota Semarang Dengan Menggunakan Sistem Informasi Geografis," *J. Geod. Undip*, vol. 8, no. 4, pp. 154–164, 2019.

- [3] R. Situmorang et al., “Model Algoritma K-Nearest Neighbor (K-NN) dan Naïve Bayes,” vol. 7, no. 1, pp. 250–254, 2023.
- [4] B. J. A. Gunadi, A. L. Nugraha, and A. Suprayogi, “Aplikasi Pemetaan Multi Risiko Bencana di Kabupaten Banyumas Menggunakan Open Source Software GIS,” *Geod. Undip Oktober*, vol. 4, no. 4, pp. 287–296, 2015.
- [5] F. Tempola, M. Muhammad, and A. Khairan, “Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, p. 577, 2018, doi: 10.25126/jtiik.201855983.
- [6] K. Pasaribu, R. Saputra, and C. Setianingsih, “Sistem Informasi Monitoring Bencana Alam Dari Data Media Sosial Menggunakan Metode K-Nearest Neighbor Natural Disaster Monitoring Information System From Social Media Data Using K-Nearest Neighbor Method,” vol. 8, no. 5, pp. 6684–6693, 2021.
- [7] S. Cumel, David Zamri, Rahmaddeni, “Perbandingan Metode Data Mining untuk Prediksi Banjir Dengan Algoritma Naïve Bayes dan KNN,” *SENTIMAS Semin. Nas. Penelit. dan ...*, pp. 40–48, 2022, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas/article/view/353%0Ahttps://journal.irpi.or.id/index.php/sentimas/article/download/353/132>.
- [8] Z. E. Fitri, L. N. Y. Syahputri, and A. M. N. Imron, “Classification of White Blood Cell Abnormalities for Early Detection of Myeloproliferative Neoplasms Syndrome Based on K-Nearest Neighbour,” *Sci. J. Informatics*, vol. 7, no. 1, pp. 136–142, 2020, doi: 10.15294/sji.v7i1.24372.
- [9] H. Azis, P. Purnawansyah, F. Fattah, and I. P. Putri, “Performa Klasifikasi K-NN dan Cross Validation Pada Data Pasien Pengidap Penyakit Jantung,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 81–86, 2020, doi: 10.33096/ilkom.v12i2.507.81-86.
- [10] R. R. R. Arisandi, B. Warsito, and A. R. Hakim, “Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation,” *J. Gaussian*, vol. 11, no. 1, pp. 130–139, 2022, doi: 10.14710/j.gauss.v11i1.33991.
- [11] N. Anggraini et al., “Analisis Pemetaan Daerah Rawan Banjir Di Kota Medan Tahun 2020,” *J. Samudra Geogr.*, vol. 4, no. 2, pp. 27–33, 2021, doi: 10.33059/jsg.v4i2.3851.
- [12] Y. Riko, A. I. Meha, and S. Y. J. Prasetyo, “Perubahan Konversi Lahan Menggunakan NDVI, EVI, SAVI dan PCA pada Citra Landsat 8 (Studi Kasus : Kota Salatiga),” *Indones. J. Comput. Model.*, vol. 1, pp. 25–30, 2019.
- [13] S. Yulianto Joko Praetyo, K. Dwi Hartomo, B. Hasiholan Simanjuntak, and D. Widiyanto Candra, “Mitigation & Identification for Local Aridity, Based of Vegetation Indices Combined with Spatial Statistics & Clustering K Means,” *J. Phys. Conf. Ser.*, vol. 1235, no. 1, 2019, doi: 10.1088/1742-6596/1235/1/012028.
- [14] P. S. Ananda, “Analisa Potensi Bencana Tanah Longsor menggunakan Interpolasi Inverse Distance Weighted (IDW),” vol. 6, no. 1, pp. 6–9, 2022.
- [15] B. Laurensz, F. Lawalata, and S. Y. J. Prasetyo, “Potensi resiko banjir dengan menggunakan citra satelit (Studi kasus: Kota Manado, Provinsi Sulawesi Utara),” *Indones. J. Comput. Model.*, vol. 2, no. 1, pp. 17–24, 2019, [Online]. Available: <https://ejournal.uksw.edu/icm/article/view/2536>.