

Desain Sistem Prediksi Wilayah Terdampak Banjir dengan Machine Learning berbasis Data Sistem Informasi Geografis

Amrul Faruq¹, Andi Syaiful Amal²

^{1,2} Program Profesi Insinyur, Universitas Muhammadiyah Malang, Jl. Raya Tlogomas 246 Malang

Kontak Person:

Amrul Faruq

Jl. Raya Tlogomas 246, Malang 65144

E-mail: faruq@umm.ac.id

Abstrak (TNR 11)

Penelitian ini bertujuan untuk merancang dan membangun sistem prediksi wilayah dampak banjir berbasis Geographic Information System (GIS) di Kota Malang. Sistem ini memanfaatkan data historis banjir, informasi topografi, dan faktor lingkungan seperti curah hujan dan tata guna lahan untuk memetakan area yang berpotensi terdampak banjir. Metode machine learning, yaitu Random Forest dan Support Vector Machine, digunakan untuk meningkatkan akurasi prediksi. Proses pengembangan sistem melibatkan pengolahan data historis dan topografi menggunakan GIS, serta penerapan algoritma prediksi berbasis machine learning. Hasil penelitian menunjukkan bahwa sistem ini mampu memprediksi wilayah rawan banjir dengan akurasi yang memadai, sehingga dapat menjadi alat pendukung bagi pemerintah daerah dan pihak terkait dalam meningkatkan kesiapsiagaan serta respons terhadap bencana banjir. Sistem prediksi ini juga diharapkan dapat mendukung perencanaan mitigasi bencana, seperti pengelolaan tata ruang yang lebih baik dan pembangunan infrastruktur yang tangguh. Dengan demikian, sistem ini berkontribusi pada pengurangan risiko bencana dan kerugian akibat banjir serta peningkatan ketahanan Kota Malang terhadap ancaman banjir di masa depan.

Kata kunci: prediksi banjir, machine learning, GIS, kecerdasan buatan

1. PENDAHULUAN

Banjir merupakan salah satu permasalahan utama yang sering terjadi di berbagai wilayah Indonesia. Sebagai negara tropis dengan intensitas hujan tinggi, Indonesia kerap menghadapi kondisi di mana air meluap dan menggenangi wilayah daratan yang seharusnya kering. Menurut Badan Nasional Penanggulangan Bencana (BNPB, 2011), banjir adalah peristiwa meluapnya air ke permukaan tanah akibat hujan deras, sungai meluap, atau drainase yang tidak memadai. Dampak banjir sangat beragam, mulai dari kerugian materi, seperti kerusakan rumah dan infrastruktur, hingga ancaman keselamatan jiwa. Kondisi ini menuntut perhatian serius, terutama di daerah yang sering mengalami kejadian serupa.

Kota Malang, Jawa Timur, menjadi salah satu daerah yang menghadapi permasalahan banjir akibat hujan dengan intensitas tinggi. Berdasarkan laporan Kepala Badan Penanggulangan Bencana Daerah (BPBD) Kota Malang, Prayitno, beberapa titik di kota ini kerap tergenang air saat musim hujan. Salah satu wilayah yang mengalami dampak parah adalah Perumahan Sigura-gura Residence di Kelurahan Karangbesuki, Kecamatan Sukun, dengan ketinggian air mencapai dua meter. Selain itu, banjir juga dilaporkan terjadi di beberapa titik lain, seperti di Kecamatan Klojen dan Kecamatan Lowokwaru. Kondisi ini menunjukkan bahwa Kota Malang memiliki kerentanan tinggi terhadap bencana banjir.

Banjir yang melanda Kota Malang membawa berbagai dampak signifikan bagi masyarakat. Selain kerugian material berupa kerusakan rumah, fasilitas umum, dan kendaraan, banjir juga mengganggu aktivitas sehari-hari, termasuk pendidikan dan ekonomi masyarakat. Tidak hanya itu, genangan air yang lama dapat meningkatkan risiko penyebaran penyakit, seperti demam berdarah dan diare. Kerugian ini menegaskan pentingnya upaya mitigasi dan penanganan banjir yang lebih baik untuk melindungi masyarakat dan meminimalkan dampak bencana.

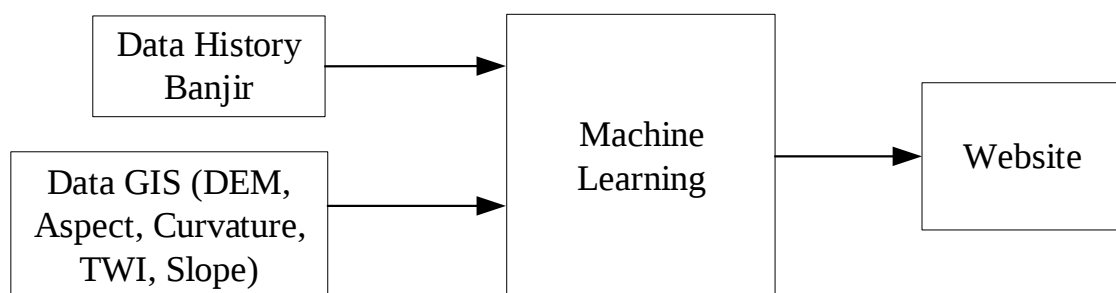
Saat ini, Kota Malang belum memiliki sistem prediksi banjir yang dapat mengidentifikasi wilayah-wilayah dengan risiko tinggi secara akurat. Padahal, informasi tersebut sangat penting untuk meningkatkan kesiapsiagaan masyarakat dan pemerintah daerah dalam menghadapi potensi banjir. Penggunaan teknologi modern, seperti Machine Learning berbasis Geographic Information System (GIS), dapat menjadi solusi efektif untuk memprediksi area terdampak. Sistem ini mampu menganalisis

berbagai faktor, seperti curah hujan, topografi, dan tata guna lahan, untuk memberikan gambaran yang lebih akurat tentang risiko banjir di masa mendatang.

Pengembangan sistem prediksi banjir berbasis GIS dengan teknologi Machine Learning akan memberikan manfaat besar, khususnya di Kota Malang. Sistem ini tidak hanya menyediakan informasi prediksi yang akurat, tetapi juga membantu pemerintah daerah dalam menyusun kebijakan mitigasi yang lebih efektif. Misalnya, informasi prediktif dapat digunakan untuk menentukan lokasi pembangunan infrastruktur anti-banjir atau perencanaan tata ruang yang lebih adaptif. Dengan adanya sistem ini, diharapkan masyarakat dan pihak terkait dapat lebih siap menghadapi ancaman banjir, sehingga dampak negatif dapat diminimalkan.

2. METODE PENELITIAN

Blok diagram perancangan sistem



Gambar 1. Blok diagram sistem yang diusulkan

1. Data History Banjir

Data History Banjir adalah data yang digunakan untuk memperkuat keakuratan prediksi suatu wilayah terdampak banjir. Data History Banjir yang didapat dari BPBD sebagai lembaga penyedia data dan sudah di preprocessing selanjutnya data tersebut di proses pada Machine Learning sebagai data yang diolah.

2. Data Sistem Informasi Geografis

Sistem Informasi Geografis (SIG) adalah suatu sistem yang dirancang untuk mengumpulkan, menyimpan, mengelola, menganalisis, dan menyajikan data yang terkait dengan lokasi geografis atau spasial. Dalam konteks SIG, data geografis seperti peta, citra satelit, dan informasi demografi menjadi elemen kunci yang diintegrasikan melalui perangkat lunak SIG, seperti ArcGIS, QGIS, dan Google Earth. Perangkat keras seperti komputer dan server digunakan untuk menyimpan, mengelola, dan mengakses data geografis ini. Proses analisis SIG melibatkan berbagai metode dan algoritma, seperti analisis jarak dan overlay peta, untuk mendapatkan informasi yang bermakna dari data spasial.

3. Machine Learning

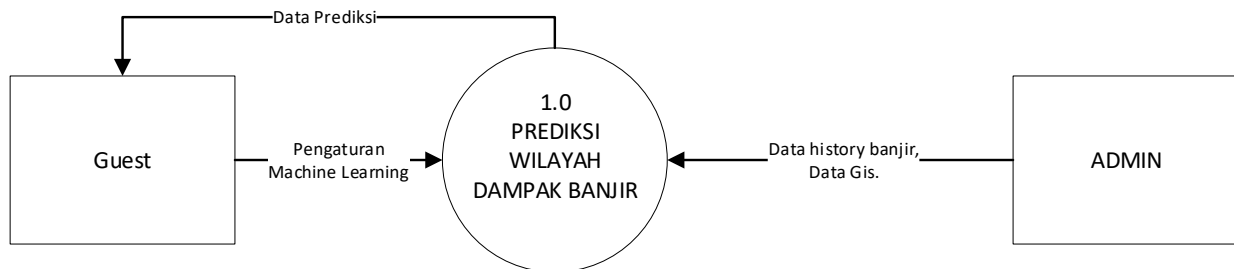
Machine learning adalah cabang AI yang paling umum digunakan. Machine learning memungkinkan mesin untuk belajar menggunakan algoritma dari data dan membuat keputusan berdasarkan pola-pola yang ditemukan dalam data tersebut.[3]

4. Website

Website adalah salah satu aplikasi yang berisikan dokumen multimedia (teks, gambar, suara, animasi, video) di dalamnya menggunakan protocol HTTP (Hypertext Transfer Protocol) dan untuk mengaksesnya menggunakan perangkat yang disebut browser.[4] Website pada Project ini digunakan sebagai media untuk menampilkan informasi dari setelah di proses pada Machine Learning. Website

yang nantinya sebagai media untuk ditampilkan pada pengguna saat menginput pembagian data training, validasi, dan test berupa peta wilayah Kota Malang.

Perancangan sistem



Gambar 2. Level diagram sistem yang diusulkan. DFD level 0 dan 1

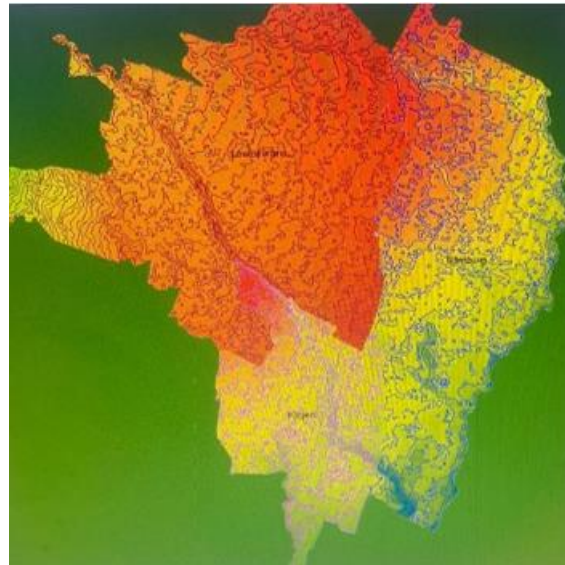
Pada DFD level 0 ini menjelaskan alur global seluruh bagian dari Sistem Prediksi Wilayah Dampak Banjir. DFD level 0 (lihat gambar 2) ini mempunyai dua external entity yaitu admin dan guest. Admin disini memasukkan data history banjir dan GIS. Dari data tersebut nantinya diolah di Machine Learning dan menghasilkan output hasil prediksi yang dapat diakses oleh guest yaitu menampilkan peta hasil prediksi bencana banjir Kota Malang.

2.3 Pendahuluan Metode

Metode sistem yang paling utama dalam pengembangan sistem ini adalah Machine Learning. Dalam prediksi wilayah dampak banjir menggunakan Random Forest dan Support Vector Machine. Pada tahap Preprocessing Data History banjir di export menjadi data shape file serta digabung dengan data tidak banjir. Data shape file tersebut diberi value dari data GIS (DEM, Aspect, Curvature, TWI, Slope) berdasarkan koordinat titik banjir dan tidak banjir. Pada tahap training data yang telah di preprocessing dibagi menjadi Data Training, Data Validasi, dan Data Test.

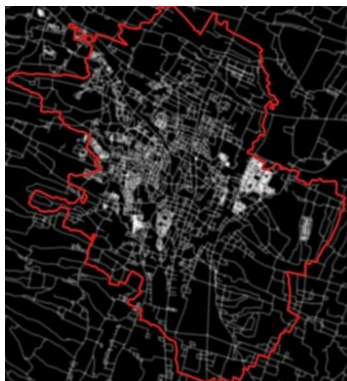
Setelah training, Random Forest dan Support Vector Machine digunakan untuk memprediksi dampak banjir khususnya Kota Malang berdasarkan input history banjir dan informasi geografis yang relevan. Penentuan metode yang diusulkan dapat dilakukan dengan mempertimbangkan beberapa faktor, seperti kebutuhan pengguna, kompleksitas sistem, waktu dan biaya pengembangan, dan kemampuan tim pengembang

Implementasi peta wilayah menggunakan aplikasi ArcGis dengan mengambil data dari aplikasi sasplanet. Implementasi counter tanah menggunakan aplikasi global mapper dengan mengambil data DEM dari website Demnas. Selanjutnya Penggabungan peta ketinggian wilayah berdasarkan MDPL menggunakan aplikasi ArcGis dengan mengolah data dari peta citra satelit dan countur tanah. Berikut Gambar 3 Implementasi peta ketinggian wilayah kota Malang berdasarkan mpdl (meter di atas permukaan laut).

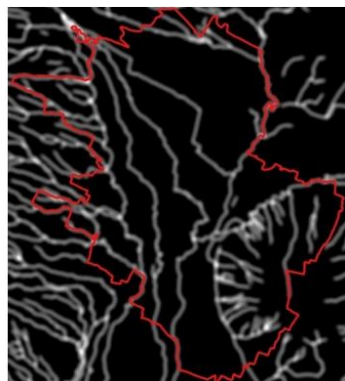


Gambar 3. Penggabungan Countur Tanah dan Peta Citra Kota Malang

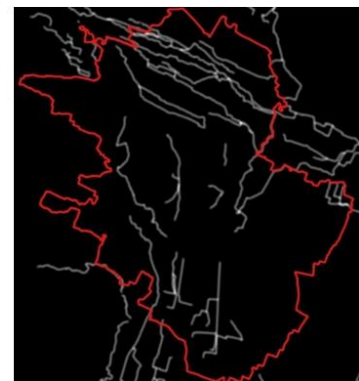
Pada gambar 8 Distance to Road digunakan untuk menampilkan jarak suatu wilayah terhadap jalan. Dimana area yang putih menggambarkan daerah yang dekat dari jalan. Pada Gambar 4 Distance to River digunakan untuk menampilkan jarak suatu wilayah terhadap Sungai. Dimana area yang putih menggambarkan daerah yang dekat dari Sungai. Pada Gambar 4 Distance to drainage digunakan untuk menampilkan jarak suatu wilayah terhadap irigasi. Dimana area yang putih menggambarkan daerah yang dekat dari irigasi.



a. distance to road



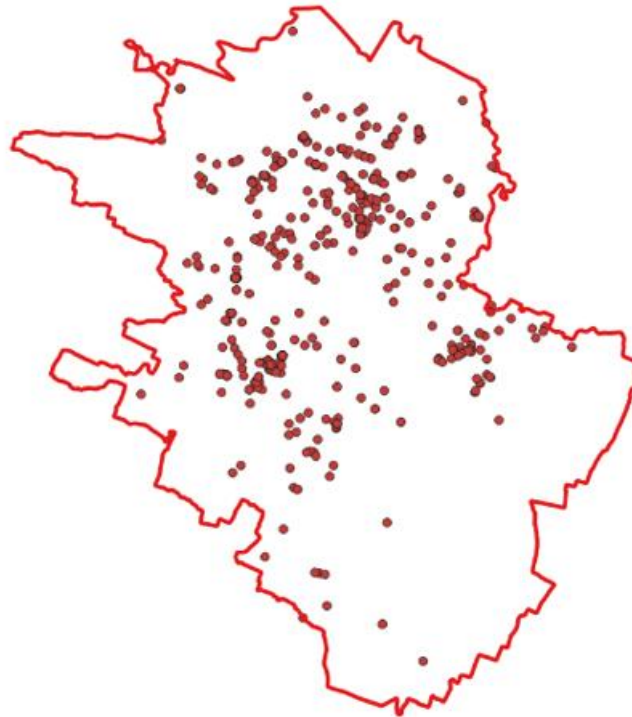
b. distance to river



c. distance to drainage

Gambar 4. Distance (jarak) peta berdasarkan jalan, sungai, dan sistem irigasi

Data history di dapatkan dari BPBD Kota Malang. Data History ini berisi data-data kejadian banjir di kota Malang yang berupa data raw. Data raw diconvert menjadi data titik menggunakan aplikasi QGIS dengan format shapefile seperti pada gambar 5 ini.



Gambar 5 Data Titik Banjir Kota Malang

Data GIS yang diperoleh terdiri dari Data DEM, Aspect, Curvature, Slope, TWI, DTRoad, DTRiver, dan DTDrainage. Semua data tersebut perlu dijadikan satu file dengan format shapefile sebelum dimasukkan ke dalam machine learning. Program yang dibuat menggunakan bahasa python untuk process preprocessing data GIS tersebut. Data DEM digunakan untuk membuat titik koordinat seluruh kota Malang. Titik koordinat tersebut kemudian digunakan sebagai patokan data value dari setiap data-data GIS yang telah di peroleh.

2.7 Machine Learning

2.7.1 Random Forest

Random forest adalah sebuah model machine learning yang digunakan untuk tugas klasifikasi dan regresi. Algoritma ini merupakan bentuk ensemble learning, yang berarti menggabungkan beberapa model prediktif yang lebih sederhana untuk mencapai prediksi yang lebih akurat dan stabil. Serta Memberikan perkiraan kepentingan fitur yang digunakan dalam prediksi.

2.7.1 Support Vector Machine

Support Vector Machine (SVM) adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. SVM memiliki prinsip dasar linier classifier yaitu kasus klasifikasi yang secara linier dapat dipisahkan, namun SVM telah dikembangkan agar dapat bekerja pada problem non-linier dengan memasukkan konsep kernel pada ruang kerja berdimensi tinggi.

2.8 Website

Aplikasi Streamlit digunakan untuk membuat tampilan website sistem yang digunakan untuk menampilkan peta hasil output prediksi bencana banjir Kota Malang. Berikut gambar 6 Tampilan Home dan gambar 7 tampilan halaman pada Machine Learning.



Gambar 6 Tampilan Home

Gambar 7 Tampilan halaman Machine Learning

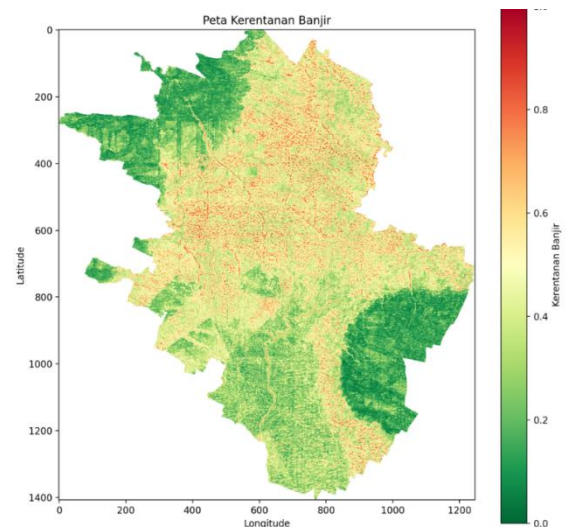
3. HASIL DAN PEMBAHASAN

3.1 Random Forest

Berdasarkan Tabel 2 Dari 8 pengujian yang dilakukan dengan pengaturan pembagian data dan jumlah pohon, dapat dilihat bahwa akurasi yang didapat berkisar 79-82%. Hal ini menunjukkan bahwa hasil pengujian memiliki akurasi yang konsisten dengan pembagian data dan jumlah pohon yang bervariasi. Peta yang dihasilkan juga memiliki pola yang dapat dikenali dimana daerah yang merah merupakan daerah dengan peluang banjir yang tinggi. Warna kuning untuk peluang sedang dan hijau untuk daerah dengan peluang banjir yang rendah.

Tabel 2 Hasil Pengujian Model Random Forest

No	Pembagian Data			Jumlah Pohon	Akurasi	Waktu
	Train	Valid	Test			
1	60%	20%	20%	100	82%	0:52
2	70%	20%	10%	100	81%	0:51
3	80%	10%	10%	100	80%	0:51
4	50%	25%	25%	100	79%	0:50
5	70%	10%	20%	100	82%	0:52
6	60%	20%	20%	200	82%	1:01
7	70%	20%	10%	200	82%	1:03
8	80%	10%	10%	200	81%	1:05
9	50%	25%	25%	200	80%	1:02
10	70%	10%	20%	200	82%	1:03
11	60%	20%	20%	300	82%	1:13
12	70%	20%	10%	300	82%	1:15
13	80%	10%	10%	300	80%	1:15
14	50%	25%	25%	300	80%	1:14
15	70%	10%	20%	300	82%	1:14



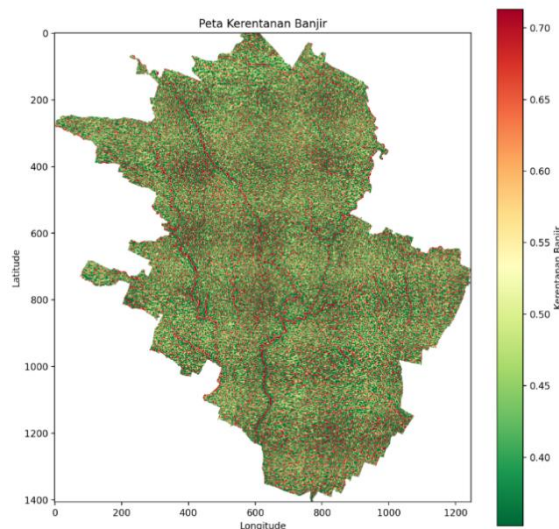
Gambar 8 Tampilan Peta Hasil Prediksi Random

3.2 Support Vector Machine

Berdasarkan Tabel 3 Dari tiga pengujian yang dilakukan dengan 3 jenis kernel yang berbeda yaitu rbf, sigmoid, dan poly. Akurasi yang didapatkan bervariasi setiap kernel, dimana kernel rbf dan poly memiliki akurasi yang mirip di angka 64-70%. Sedangkan, kernel sigmoid memiliki akurasi yang rendah dari kernel lain di angka 49%-58%. Setelah dievaluasi ketiga kernel tersebut, kernel rbf yang memiliki pola peta kerentanan yang paling jelas dibandingkan kernel yang lain. Kernel rbf menampilkan peta yang ideal dengan 3 warna klasifikasi kerentanan banjir yang ditetapkan. Sedangkan kernel sigmoid dan poly menampilkan peta yang didominasi oleh warna kuning (kerentanan banjir sedang).

Tabel 3 Pengujian Support Vector Machine

No	Pembagian Data			Jenis Kernel	Akurasi	Waktu
	Train	Valid	Test			
1	60%	20%	20%	rbf	69%	1:42
2	70%	20%	10%	rbf	64%	1:53
3	70%	10%	20%	rbf	69%	1:53
4	60%	20%	20%	sigmoid	49%	1:30
5	70%	20%	10%	sigmoid	58%	1:29
6	70%	10%	20%	sigmoid	50%	1:33
7	60%	20%	20%	poly	70%	1:14
8	70%	20%	10%	poly	64%	1:17
9	70%	10%	20%	poly	69%	1:18



Gambar 9 Tampilan Hasil Prediksi Support Vector Machine rbf

4. KESIMPULAN

Berdasarkan hasil pengujian yang dilakukan yang telah dilakukan, maka dapat diambil beberapa kesimpulan, diantaranya:

1. Berdasarkan hasil pengujian Random Forest dengan pembagian data 80% train, 10% validasi, dan 20% test serta 200 pohon memberikan akurasi tertinggi sebesar 83%. Secara umum, akurasi berkisar antara 78% hingga 83% tergantung pada konfigurasi pohon dan pembagian data.
2. Pengujian model SVM juga dilakukan dengan berbagai skenario pengaturan jenis kernel dan pembagian persentase data. Kernel yang diuji adalah RBF, sigmoid, dan poly. Hasil terbaik diperoleh dengan kernel poly yang mencapai akurasi sebesar 70%, sementara kernel RBF menghasilkan akurasi 68% dan kernel sigmoid memiliki akurasi terendah yaitu 49%
3. Pengujian website dilakukan untuk memastikan semua fungsi dasar beroperasi dengan baik, seperti menampilkan halaman, navigasi, dan masukan data dari pengguna. Semua pengujian berhasil dan dinyatakan valid, menunjukkan bahwa website berfungsi sesuai harapan dan mampu menampilkan hasil dari model machine learning dengan baik.
4. Random Forest memberikan performa yang lebih baik dan konsisten dibandingkan SVM, dengan akurasi tertinggi 83% versus 70% pada SVM. Random Forest juga menunjukkan stabilitas hasil yang lebih tinggi dengan variasi akurasi yang lebih kecil.

REFERENSI

- [1] Komang Ika Hendriana, I Gede Adi Saputra Yasa, Made Windu Antara Kesiman, I Made Gede Sunarya, Juli 2013, Sistem Informasi Geografis Penentuan wilayah Rawan Banjir di kabupaten Buleleng, Singaraja, Bali, KARMAPATI (Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika), Vol. 2, No. 5.
- [2] suarasurabaya.net, Hujan Deras Sebabkan Banjir di Sejumlah Titik di Kota Malang, <https://www.suarasurabaya.net/kelanakota/2023/hujan-deras-sebabkan-banjir-di-sejumlah-titik-di-kota-malang/> diakses pada tanggal 27 November 2023
- [3] Dimas Taufiq, Dwi Duta M., Faizal M, Muhamad Ghifari A., Muhammad Miftahur R., Risa Azzahra Z., Octaviano Eka S., Rully Mujiastuti, Nurvelly R., Yana A., Oktober 2023, Edukasi dan Pelatihan Pengenalan Machine Learning dan Computer Vision Untuk Mengeksplorasi Potensi Visual, Jakarta, Universitas Muhammadiyah Jakarta, Seminar Nasional LPPM UMJ.
- [4] Tuti Susilawati, Fanny Yuliansyah, Muhammad Romzi, Rintan Aryani, Juni 2020, Membangun Website Toko Online Pempek NTHREE Menggunakan PHP dan MySQL, Universitas Mahakarya Asia, Jurnal Teknik Informatika Mahakarya, Vol.3, No. 1.
- [5] Omar Seleem, Georgy Ayzel, Arthur Costa Tomaz de Souza, Axel Bronstert & Maik Heistermann (2022) Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany, *Geomatics, Natural Hazards and Risk*, 13:1, 1640-1662, DOI: 10.1080/19475705.2022.2097131.
- [6] Rona Nisa Sofia Amriza, Didi Supriyadi, Oktober 2021, Komparasi Metode Machine Learning dan Deep Learning untuk Deteksi Emosi pada Text di Sosial Media, Purwokerto, Institut Teknologi TelkomPurwokerto, jurnal JUPITER Vol. 13 No.2.