

Klasifikasi Level Banjir Menggunakan Random Forest dan Support Vector Machine

Larasati Syarafina Qamarani¹, Mardhani Riasetiawan^{*2}

^{1,2}Departemen Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta, Indonesia

e-mail: ¹larasati.syarafina@mail.ugm.ac.id, ^{*2}mardhani@ugm.ac.id

Abstrak

Banjir merupakan salah satu bencana alam yang umum terjadi di Indonesia. Semakin tinggi tingkat banjir, semakin besar dampak yang diakibatkan oleh bencana ini. Mulai dari kerugian pada aspek ekonomi, lingkungan, infrastruktur, hingga korban jiwa. Rekapitulasi dampak kerugian dari setiap peristiwa banjir seperti durasi, ketinggian air, dan jumlah terdampak dianalisis untuk mengidentifikasi karakteristik banjir sesuai dengan tingkatan banjir yang terjadi. Data iklim seperti temperatur, kelembaban, curah hujan, dan kecepatan angin diteliti sebagai parameter ciri terjadinya banjir. Penelitian ini bertujuan untuk memperoleh hasil klasifikasi level banjir dengan tolak ukur hasil sistem klasifikasi menggunakan Random Forest dan Support Vector Machine serta evaluasi akurasi dengan Confusion Matrix yang diharapkan dapat menjadi pertimbangan untuk mengambil keputusan saat terjadi banjir dalam upaya meminimalisir kerugian akibat banjir. Proses penelitian menggunakan data historis banjir dari BPBD DKI Jakarta melalui website Satu Data Jakarta serta data iklim dari Stasiun Geofisika BMKG yang diambil dalam rentang tahun 2013-2020. Sistem klasifikasi Random Forest mencapai hasil sangat tinggi dengan akurasi sangat baik yaitu sebesar 99,21%. Sementara sistem klasifikasi Support Vector Machine mencapai hasil sedikit lebih rendah di angka 98,43%. Kedua sistem mengalami overfitting di awal pembuatan model, namun semakin menurun seiring bertambahnya data penelitian sehingga kemampuan generalisasi meningkat.

Kata kunci— Banjir, Dampak, Level, Random Forest, Support Vector Machine

Abstract

Floods are one of the most common natural disasters in Indonesia, with higher flood levels correlating with greater impacts, including economic losses, environmental degradation, infrastructure damage, and human casualties. This study analyzes the impact of each flood event by examining factors such as duration, water level, and the number of affected individuals to identify flood characteristics based on severity. Climate variables such as temperature, humidity, rainfall, and wind speed were investigated as parameters characterizing flood occurrences. The primary objective of this research is to classify flood levels using Random Forest and Support Vector Machine (SVM) algorithms, and to evaluate the accuracy of these classifications using a Confusion Matrix. The outcomes are intended to inform decision-making processes during floods, thereby aiming to minimize associated losses. The research utilized historical flood data from the DKI Jakarta BPBD, accessed through the Satu Data Jakarta website, and climate data from the BMKG Geophysical Station, covering the period from 2013 to 2020. The Random Forest classification system demonstrated exceptional performance, achieving an accuracy of 99.21%. Similarly, the SVM classification system performed robustly, with an accuracy of 98.43%. Both models initially exhibited overfitting during the early stages of model development. However, this issue is diminished as the dataset size increases, thereby enhancing the models' generalization capabilities.

Keywords— Flood, Impacts, Level, Random Forest, Support Vector Machine

1. PENDAHULUAN

Indonesia merupakan negara yang memiliki risiko besar mengalami terjadinya bencana alam. Menurut Badan Nasional Penanggulangan Bencana (BNPB), pada tahun 2022, banjir merupakan bencana alam yang paling sering terjadi di Indonesia, dengan persentase sebesar 43% dari total bencana yang terjadi di Indonesia. Bencana ini menyebabkan kerusakan yang signifikan pada infrastruktur, pertanian, ekonomi, serta dapat mengakibatkan jatuhnya korban jiwa [1].

Semakin tinggi besaran banjir yang terjadi, dampak yang diakibatkan bencana ini dapat mempengaruhi segala aspek. Bencana ini dipengaruhi oleh perubahan iklim, dimana terjadi perubahan curah hujan yang lebih intens [2]. Perubahan iklim global yang meningkat seiring berjalannya waktu mempengaruhi kelembaban, kecepatan angin, dan suhu sehingga menjadi ciri terjadinya peristiwa banjir. faktor lain seperti badai, tanah longsor, dan kerusakan bendungan dapat menyebabkan terjadinya banjir [3]. Untuk mengatasi bencana ini, para peneliti berusaha membuat suatu sistem deteksi banjir. Namun, kurangnya data yang akurat dan aplikasi dalam kejadian nyata secara tepat waktu menyebabkan sistem yang ada belum dapat meminimalisir dampak yang terjadi dalam suatu upaya mitigasi banjir [4].

Penelitian yang sudah dilakukan sebelumnya dengan fokus terhadap pengembangan sistem peringatan dini banjir banyak menggunakan algoritma Machine Learning, terutama dengan arsitektur klasifikasi seperti Random Forest dan Support Vector Machine [5]. Algoritma klasifikasi ini telah terbukti dapat bekerja efektif dalam memprediksi banjir dan mengidentifikasi area yang berisiko terkena banjir. Namun, penelitian yang tersedia terbatas pada wilayah tertentu dan belum berpengaruh pada manajemen risiko banjir di wilayah Indonesia.

Penelitian ini bertujuan untuk mengembangkan sistem berbasis Machine Learning untuk mengklasifikasikan tingkat banjir berdasarkan dampak yang terjadi beserta karakteristik cuaca pada setiap kejadian banjir di DKI Jakarta. Data historis banjir direkapitulasi oleh BPBD DKI Jakarta dapat diakses pada platform Satu Data Jakarta (<https://satudata.jakarta.go.id/open-data/>). Data iklim harian diambil dengan parameter yang dibutuhkan dalam klasifikasi melalui Data Online BMKG (<https://dataonline.BMKG.go.id>), yaitu parameter berupa curah hujan, temperatur, kelembaban, dan kecepatan angin. Penelitian ini akan berfokus untuk membuat suatu model yang dapat mengklasifikasikan seluruh data karakteristik banjir menggunakan algoritma Random Forest dan Support Vector Machine.

2. METODE PENELITIAN

2.1 Pengumpulan Data

Pembuatan sistem klasifikasi berbasis algoritma Random Forest dan Support Vector Machine meliputi beberapa tahapan, antara lain perancangan, persiapan data, pelatihan, pengujian, dan evaluasi sistem klasifikasi. Data historis banjir diperoleh BPBD Jakarta mencakup rekapitulasi lama genangan, ketinggian air, jumlah pengungsi, jumlah wilayah terdampak, periode kejadian, lokasi kejadian, dan jumlah korban terdampak, yang dilengkapi dengan label kelas banjir. Data iklim diperoleh dari Stasiun Klimatologi BMKG secara harian selama 8 tahun dengan parameter berupa kecepatan angin, curah hujan, suhu, dan kelembaban.

Data historis banjir dan data iklim digunakan sebagai data pelatihan dan pengujian pada pembuatan model. pengumpulan dataset meliputi pengunduhan data dari kedua sumber dan

penggabungan data BPBD dan data BMKG berdasarkan area dan periode yang sama untuk setiap kejadian banjir. Data iklim dari stasiun BMKG dikumpulkan berdasarkan setiap peristiwa banjir yang terjadi di DKI Jakarta yang direkapitulasi BPBD selama tahun 2013-2020. Level banjir yang tersedia dalam dataset dan ditunjukkan pada Tabel 1 telah sesuai dengan standar Modul Sistem Pengendalian Banjir dari Kementerian Pekerjaan Umum dan Perumahan Rakyat di Indonesia.

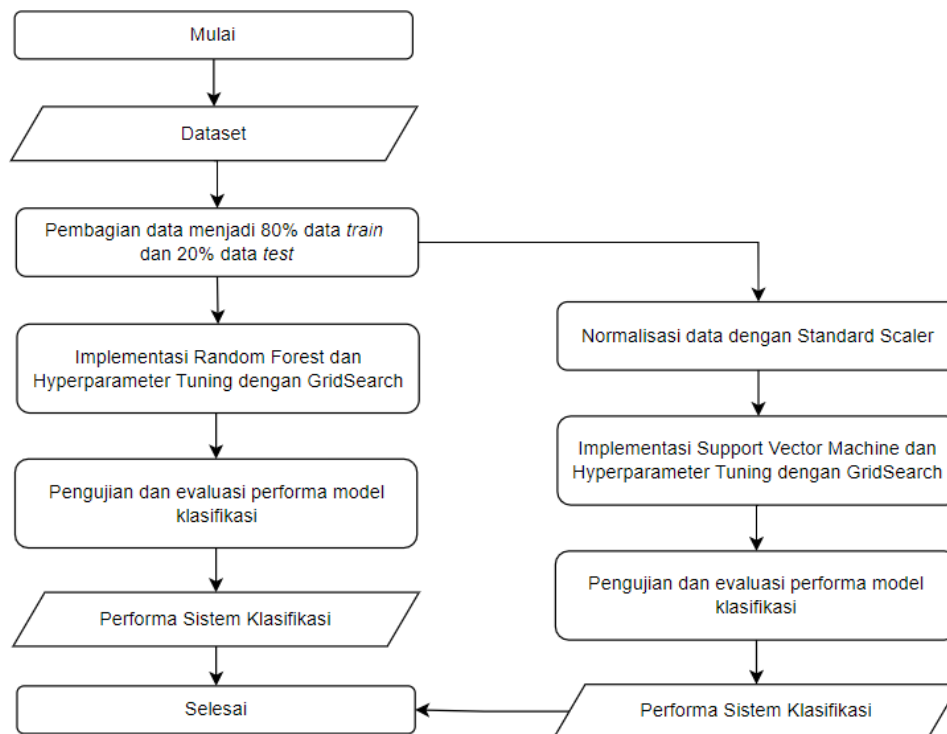
Tabel I Flood Level

No.	Level Banjir	Tingkat Siaga	Ketinggian Air
1	Rendah	Siaga 3	< 50 cm
2	Sedang	Siaga 2	50 - 125 cm
3	Tinggi	Siaga 1	> 125 cm

Dataset ini terdiri dari 633 kejadian banjir yang akan dibagi menjadi data pelatihan dan data pengujian dalam rasio 4:1, sehingga 80% data akan digunakan dalam proses pelatihan. Sedangkan 20% data yang tidak digunakan pada proses pelatihan akan digunakan untuk pengujian model dan akan dievaluasi menggunakan konsep confusion matrix.

2.2 Implementasi Model Klasifikasi

Setelah pengumpulan data dari kedua sumber, data historis banjir dan data iklim akan digabungkan untuk dilakukan pembersihan data. Proses ini terdiri dari pembersihan data yang tidak diperlukan dalam model klasifikasi dan memanipulasi data yang bernilai kosong atau data yang tidak terukur menggunakan nilai mean atau median berdasarkan distribusinya. Feature encoding akan dilakukan pada fitur yang digunakan sebagai label target, yaitu fitur Banjir, kolom yang merepresentasikan tingkat banjir dari setiap kejadian. Nilai pada fitur tersebut akan dikonversikan dalam bentuk numerik dengan Rendah sebagai 0, Sedang sebagai 1, dan Tinggi sebagai 2. Selanjutnya, data akan diimpor ke dalam model untuk diklasifikasikan ke dalam tiga kelas level banjir seperti yang ditunjukkan dalam Gambar 1.



Gambar 1 Diagram Alir Implementasi Sistem Klasifikasi

Model klasifikasi akan dibuat menjadi dua model yang berbeda dengan algoritma yang berbeda. Setiap model akan menjalankan hyperparameter tuning menggunakan Grid Search dengan cross validation untuk menemukan kombinasi parameter terbaik. Pada model Random Forest, setelah mendefinisikan data yang sudah terpisah, modul Random Forest akan diinisialisasi terlebih dahulu. Proses hyperparameter tuning akan mencari kombinasi nilai terbaik dari parameter yang dibutuhkan dalam arsitektur Random Forest yaitu $n_estimator$, max_depth , $min_samples_split$, dan $min_samples_leaf$. Pada model Support Vector Machine, data yang telah diimpor ke dalam model dinormalisasi terlebih dahulu menggunakan Standard Scaler. Hyperparameter tuning akan dilakukan untuk menemukan kombinasi parameter C dan tipe kernel yang dapat mencapai hasil terbaik. Dalam tahap ini, kedua model akan menemukan model dengan kombinasi terbaik berdasarkan hasil hyperparameter tuning.

2.3 Pengujian dan Evaluasi

Pada tahapan pengujian dan evaluasi, untuk melihat seberapa baik model bekerja pada data pelatihan, kurva pembelajaran digunakan untuk menunjukkan bagaimana kinerja model pada data pelatihan dan pengujian seiring bertambahnya ukuran set data atau jumlah iterasi pelatihan. Kurva ini akan membantu mengidentifikasi apakah model mengalami overfitting atau underfitting. Proses pengujian pada sistem akan dilakukan dengan mengklasifikasikan seluruh data pengujian ke dalam tiga kelas banjir pada masing-masing model klasifikasi. Evaluasi akan dilakukan dengan memanfaatkan pengukuran dengan konsep perhitungan Confusion Matrix. Konsep ini akan membandingkan keluaran model dengan label data aktual, sehingga menghasilkan nilai metrik akurasi, presisi, recall, dan f1-score dari masing-masing model sesuai dengan persamaan (1) hingga persamaan (4).

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Presisi = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F\text{-score} = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

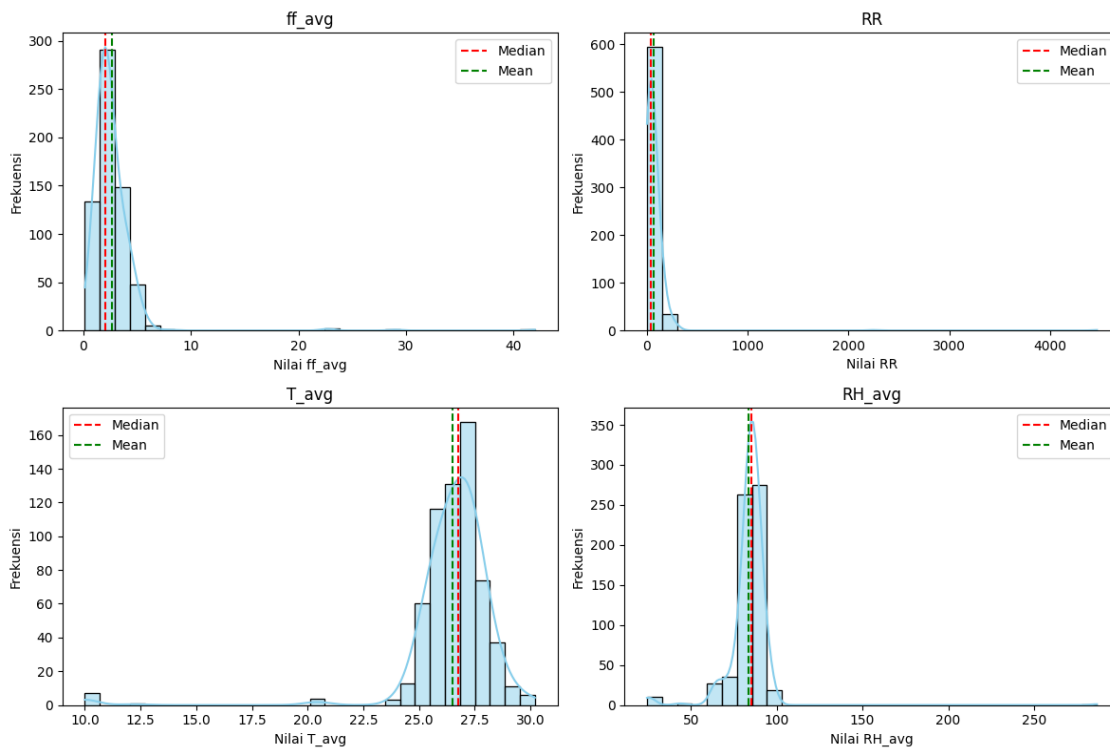
3. HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini, yaitu data historis banjir dari BPBD dan data iklim dari BMKG, dibagi menjadi data latih dan data uji. Distribusi data pada masing-masing kelas mengalami ketidakmerataan, namun pada penelitian ini tidak diberikan perlakuan pada ketidakseimbangan data untuk menjaga keaslian data dan melatih model agar dapat mempelajari data seperti pada kejadian di dunia nyata. Dalam data historis banjir, terdapat data yang merepresentasikan wilayah dan waktu kejadian. Data ini akan dihapus karena tidak digunakan dalam klasifikasi level banjir.

Setelah membersihkan data yang tidak diperlukan, manipulasi data dilakukan untuk mengisi data yang bernilai 0, nol, dan tidak terukur. Variabel yang memiliki data seperti ini adalah $Tavg$, RH_avg , RR , dan ff_avg . Merujuk pada Gambar 2, fitur ff_avg dan RR menunjukkan kemiringan "skewed negative", yaitu ketika nilai rata-rata lebih rendah dari nilai

median data. Sedangkan fitur RH_avg dan T_avg menunjukkan kemiringan "skewed positive", ketika nilai median data memiliki nilai yang lebih rendah dari nilai rata-rata data.

Dengan adanya kondisi ini, manipulasi data dilakukan sesuai dengan karakteristik dari masing-masing fitur. Nilai-nilai yang dimanipulasi pada fitur dengan kemiringan "skewed negative" akan digantikan menggunakan nilai median, sedangkan nilai rata-rata akan diberikan pada nilai-nilai pada fitur dengan kemiringan "skewed positive". Hal ini bertujuan untuk mengarahkan distribusi ke arah distribusi normal.



Gambar 2 Distribusi Fitur Data

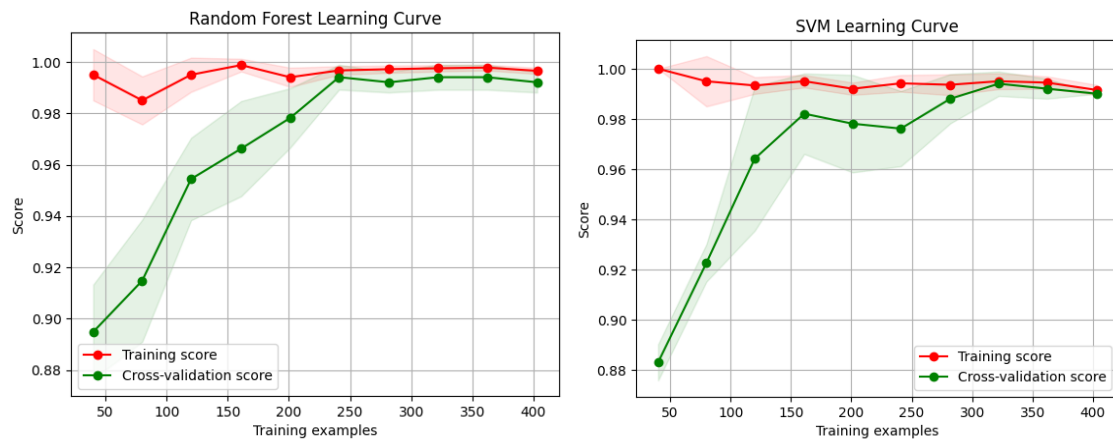
Setelah melakukan manipulasi data pada proses pra-pemrosesan data, dilakukan feature encoding untuk mengubah data pada kolom Banjir menjadi data numerik sehingga dapat digunakan sebagai label target pada sistem klasifikasi. Proses ini akan memetakan kelas-kelas banjir sesuai dengan tingkatannya, yaitu Rendah (0), Sedang (1), dan Tinggi (2). Data yang telah diproses akan dibagi menjadi data latih dan data uji dengan pembagian seperti yang ditunjukkan pada Tabel 2.

Tabel 2 Distribusi Kelas Banjir dalam dataset

No.	Data	Level Banjir	Jumlah Kejadian
1	Data Pelatihan	Rendah (Low)	251
2		Sedang (Medium)	70
3		Tinggi (Tinggi)	183
4	Data Pengujian	Rendah (Low)	60
5		Sedang (Medium)	16
6		Tinggi (Tinggi)	51

Pada tahap implementasi model klasifikasi, modul Random Forest dan Support Vector Machine (SVM) diinisialisasi di awal pembuatan masing-masing model. Kedua model akan di-tuning untuk menemukan parameter terbaik dengan menggunakan Grid Search berdasarkan hasil cross validation. Model Random Forest mencapai performa terbaik dengan menggunakan 50 n estimator, tanpa max dept, 10 min samples split, dan 1 min samples leaf sebagai parameter model, dengan nilai cross validation sebesar 99,21%. Model SVM mencapai hasil terbaik menggunakan 10 parameter regularisasi, atau C, dan tipe kernel linear, dengan nilai cross validation sebesar 98,81%.

Model klasifikasi dengan parameter yang telah ditentukan akan diuji pada data uji dan dievaluasi untuk mengetahui dan menilai kinerja performa dari masing-masing sistem. Dalam proses evaluasi kinerja model, digunakan kurva pembelajaran untuk menilai perkembangan model berdasarkan metrik akurasi yang ditunjukkan pada Gambar. 3. Kedua model mencapai akurasi yang hampir sempurna selama pelatihan. Hasil cross validation dari kedua model dimulai dari angka yang tinggi, namun jika dibandingkan dengan skor saat pelatihan model, skor cross validation menunjukkan nilai yang rendah. Seiring bertambahnya data, skor ini meningkat dengan cepat data dan hampir menyamai skor pelatihan.



Gambar 3 Perbandingan learning curve pada model klasifikasi Random Forest dan SVM

Dengan mengimplementasikan model pada 20% data yang belum pernah dilihat pada pelatihan yakni data uji, hasil perhitungan confusion matrix untuk masing-masing model dapat dilihat pada Tabel 3. Metode confusion matrix ini menunjukkan nilai true positive (TP), true negative (TN), false positive (FP), dan false negative (FN), dimana nilai-nilai tersebut menggambarkan seberapa baik model dapat mengklasifikasikan data ke dalam kelas yang benar.

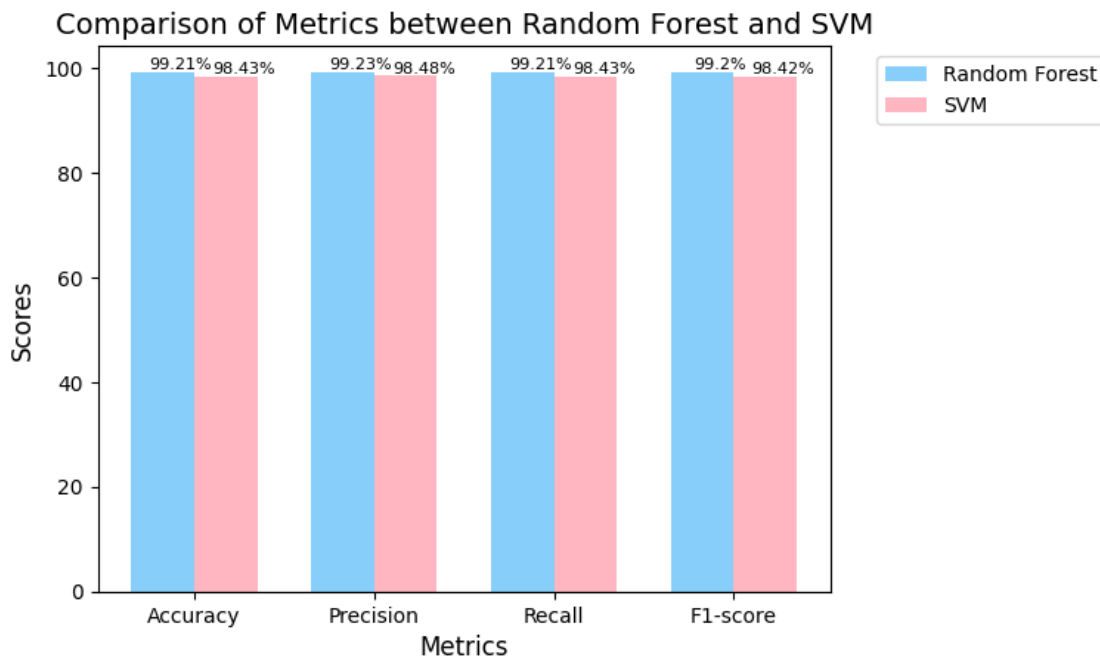
Tabel 3 Perbandingan hasil confusion matrix pada model klasifikasi Random Forest dan SVM

Model	Level	TP	TN	FP	FN
Random Forest	Rendah	51	76	0	0
	Sedang	60	66	1	0
	Tinggi	15	111	0	1
Support Vector Machine	Rendah	50	76	0	1
	Sedang	60	65	1	0
	Tinggi	15	111	0	1

Berdasarkan hasil pada Tabel 3, kedua model memiliki kemampuan generalisasi baik dengan kesalahan prediksi yang minim. Dari nilai-nilai tersebut, masing-masing model klasifikasi dievaluasi dengan memanfaatkan teori confusion matrix berdasarkan rumus pada Persamaan (1) hingga Persamaan (4), sehingga menghasilkan nilai akurasi, presisi, recall, dan f1-score yang ditunjukkan pada Tabel 4 serta grafik perbandingan pada Gambar 4.

Tabel 4 Perbandingan evaluasi model klasifikasi Random Forest dan SVM

Model	Akurasi	Presisi	Recall	F1-Score
Random Forest	0.9921	0.9923	0.9921	0.99210
Support Vector Machine	0.9843	0.9848	0.9843	0.9842



Gambar 4 Perbandingan metrik evaluasi model klasifikasi Random Forest dan SVM

Mengacu pada Gambar 4, metrik evaluasi antara model Random Forest dan SVM menunjukkan hasil yang sangat tinggi pada masing-masing model. Kedua model mencapai hasil yang hampir sempurna, dengan model Random Forest menghasilkan nilai yang sedikit lebih tinggi yaitu 99% dan SVM mencapai nilai 98%. Hasil metrik presisi, recall, dan f1-score pada masing-masing model juga memiliki nilai yang hampir sama. Hal ini menunjukkan bahwa kedua model memiliki performa yang baik dalam mengklasifikasikan data banjir ke dalam masing-masing kelas, yang direpresentasikan dengan tingginya nilai akurasi.

4. KESIMPULAN

Data historis banjir yang meliputi lama genangan, ketinggian air, jumlah pengungsi, jumlah wilayah terdampak, dan jumlah korban terdampak, dilengkapi dengan data iklim berupa temperatur, kelembapan, curah hujan, dan kecepatan angin, dapat digunakan untuk mengklasifikasikan kejadian banjir berdasarkan tingkatannya. Model Random Forest dengan parameter berupa 50 n estimator, no max depth, 10 min samples split, dan 1 min samples leaf

menghasilkan akurasi terbaik sebesar 99,21%. Model SVM dengan 10 parameter regularisasi, atau C dan kernel linear menghasilkan nilai yang sedikit di bawah Random Forest, tetapi tetap bernilai sangat tinggi, yaitu akurasi di angka 98,43%.

Pada saat pengujian model, terlihat adanya overfitting terutama pada awal proses. Kondisi ini terlihat dari perbandingan hasil learning curve pada saat proses training dan validasi, namun kondisi overfitting ini menurun seiring dengan bertambahnya iterasi data. Peningkatan data dianggap membantu model dalam mempelajari data dengan lebih baik dan meningkatkan performa klasifikasi. Berdasarkan hasil confusion matrix masing-masing model klasifikasi, kedua model dapat mengklasifikasikan kelas "Rendah" dan "Sedang" dengan performa hampir sempurna, sehingga menghasilkan nilai precision dan recall yang tinggi. Sedangkan untuk kelas "Tinggi", terdapat kesalahan prediksi yang minim. Hal ini mengindikasikan bahwa model-model tersebut memiliki kinerja yang sangat baik dalam mengklasifikasikan kelas mayoritas, tetapi masih diperlukan pelatihan lebih lanjut untuk mengidentifikasi kelas minoritas dengan lebih baik. Model Random Forest dan SVM memiliki performa yang sangat baik untuk mengklasifikasikan tingkat banjir pada data ini, yang ditunjukkan dengan hasil performa yang sangat tinggi dan kemampuan generalisasi yang baik

5. SARAN

Untuk meningkatkan penelitian ini agar dapat dipergunakan dalam kondisi nyata dan memberikan kontribusi di lapangan, saran untuk penelitian selanjutnya adalah dengan menambahkan variabel lain untuk melengkapi karakteristik banjir, seperti elevasi tanah, jenis tanah pada daerah yang diteliti, dan data hidrologi. Data tersebut dapat digunakan untuk mengenal karakteristik banjir secara lebih mendalam dan menambah keunikan data pada masing-masing kelas. Penggunaan data yang lebih banyak dapat mendukung sistem klasifikasi yang lebih baik, meminimalisir overfitting, dan meningkatkan kemampuan generalisasi, terutama ketika model tersebut digunakan untuk keperluan di dunia nyata. Selain itu, dapat dibuat suatu sistem prediksi yang memanfaatkan klasifikasi tingkat banjir ini sehingga sistem dapat bekerja secara praktis dan dapat langsung dimanfaatkan di dunia nyata secara real-time.

DAFTAR PUSTAKA

- [1] Kundzewicz, Z.W., Pińskwar, I. and Brakenridge, G.R., 2018. Changes in river flood hazard in Europe: a review. *Hydrology research*, 49(2), pp.294-302, 2017 [Online]. Available:<https://iwaponline.com/hr/article/49/2/294/37824/Changes-in-river-flood-hazard-in-Europe-a-review>. [Accessed: 19-June-2023]
- [2] Behera, J., 2020. Classifying Flood Severity Using Machine Learning. *Doctoral dissertation, Dublin, National College of Ireland*, Nov. 2021 [Online]. Available: https://www.researchgate.net/publication/356372808_A_machine_learning_approach_to_flood_severity_classification_and_alerting. [Accessed: 17-June-2023]
- [3] Sharma, P., Kar, B., Wang, J. and Bausch, D., "A Machine Learning approach to flood severity classification and alerting," *ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities*, pp. 42-47, Nov. 2021 [Online]. Available:

- <https://www.semanticscholar.org/paper/A-Machine-Learning-Approach-to-Flood-Depth-and-1A-B-Tiampo-Woods/1b085fdc46e5cb386e4f961c024129c39d89e656>. [Accessed: 16-June-2023]
- [4] Alipour, A., Ahmadalipour, A., Abbaszadeh, P. and Moradkhani, H., 2020. Leveraging Machine Learning for predicting flash flood damage in the Southeast US. *Environmental Research Letters*, 15(2), p.024011, Feb. 2020 [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-9326/ab6edd>. [Accessed: 20-June-2023]
- [5] Khan, T.A., Shahid, Z., Alam, M., Su'ud, M.M. and Kadir, K., 2019, December. Early flood risk assessment using machine learning: A Comparative Study of SVM, Q-SVM, KNN, and LDA. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, pp. 1-7, 2019 [Online]. Available: <https://ieeexplore.ieee.org/document/9024796>. [Accessed: 16-June-2023]
- [6] Khalaf, M., Hussain, A.J., Al-Jumeily, D., Baker, T., Keight, R., Lisboa, P., Fergus, P. and Al Kafri, A.S., 2018, July. A data science methodology based on Machine Learning algorithms for flood severity prediction. *IEEE Congress on Evolutionary Computation (CEC)*, pp. 1-8, Jul. 2020 [Online]. Available: <https://jurnal.ugm.ac.id/ijccs/article/view/11187>. [Accessed: 17-June-2023]
- [7] Fitriyaningsih, I., Basani, Y. and Ginting, L.M., Machine Learning: Prosperity Of Rainfall, Water Discharge, And Flood With Web Application In Deli Serdang-Aplikasi Prediksi Curah Hujan, Debit Air, dan Kejadian Banjir Berbasis Web dengan Machine Learning di Deli Serdang. *Jurnal Penelitian Komunikasi dan Opini Publik*, 22(2), p.272740, Dec. 2018 [Online]. Available: <https://www.neliti.com/publications/272740/machine-learning-prosperity-of-rainfall-water-discharge-and-flood-with-web-appli>. [Accessed: 1-July-2023]
- [8] Grady, F., Tarigan, J.K., Wahidiyat, J.R. and Prasetyo, A., November. Classification of Flood Alert in Jakarta with Random Forest. *2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA)*, pp. 1-6, 2022 [Online]. Available: <https://ieeexplore.ieee.org/document/9971411>. [Accessed: 22-June-2023]
- [9] Kim, D., Park, J., Han, H., Lee, H., Kim, H.S. and Kim, S., 2023. Application of AI-Based Models for Flood Water Level Forecasting and Flood Risk Classification. *KSCE Journal of Civil Engineering*, pp.1-12, May 2023 [Online]. Available: <https://link.springer.com/article/10.1007/s12205-023-2175-5>. [Accessed: 1-February-2024]
- [10] Sharma, T., Pal, A., Kaushik, A., Yadav, A. and Chitragupta, A., 2022, February. A Survey on Flood Prediction analysis based on ML Algorithm using Data Science Methodology. *2022 IEEE Delhi Section Conference (DELCON)* pp. 1-8, 2022 [Online]. Available: <https://ieeexplore.ieee.org/document/9753396>. [Accessed: 16-Feb-2017]