Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

SISTEM PERINGATAN DINI BANJIR BERBASIS SUPPORT VECTOR MACHINE YANG DIOPTIMALKAN DENGAN PARTICLE SWARM OPTIMIZATION

Muhammad Rizal H[1*], Fitriana M Sabir[2], Mursalim[3]

Program Studi Teknik Informatika[1,3]
Program Studi Teknologi Informasi[2]
Universitas Teknologi Akba Makassar

*Correspondent E-mail: rizal@unitama.ac.id

Authors Email: rizal@unitama.ac.id[1], fitriana@unitama.ac.id[2], mursalim@unitama.ac.id[3]

Received: December 12,2025**. Revised:** January 25,2026**. Accepted:** January 27, 2026**. Issue Period:** Vol.10 No.1 (2026), Pp. 155-167

Abstrak: Banjir merupakan salah satu bencana hidrometeorologi yang paling sering terjadi dan menimbulkan dampak besar terhadap keselamatan manusia, kerusakan infrastruktur, serta kerugian ekonomi. Kondisi ini menuntut pengembangan sistem peringatan dini banjir yang akurat, andal, dan mampu memberikan informasi secara tepat waktu. Penelitian ini bertujuan untuk mengembangkan sistem peringatan dini banjir berbasis Support Vector Machine yang dioptimalkan menggunakan Particle Swarm Optimization guna meningkatkan kinerja prediksi kejadian banjir. Metode yang digunakan meliputi pembangunan model Support Vector Machine dengan proses optimasi hiperparameter menggunakan Particle Swarm Optimization. Kinerja model yang diusulkan dibandingkan dengan Support Vector Machine tanpa optimasi dan Support Vector Machine dengan metode GridSearch. Evaluasi performa dilakukan menggunakan berbagai metrik, yaitu akurasi, presisi, recall, F1-score, spesifisitas, serta dianalisis melalui confusion matrix, kurva pembelajaran, kurva Receiver Operating Characteristic, dan konvergensi optimasi. Hasil penelitian menunjukkan bahwa model Support Vector Machine Particle Swarm Optimization memberikan performa paling unggul dengan akurasi sekitar 96% dan F1-score optimal sebesar 0,9777. Model ini mampu mendeteksi seluruh kejadian banjir tanpa kesalahan false negative dan hanya menghasilkan satu false positive, yang menunjukkan tingkat sensitivitas dan keandalan yang sangat tinggi. Analisis kurva pembelajaran memperlihatkan kemampuan generalisasi yang stabil, sedangkan proses optimasi menunjukkan konvergensi yang cepat dan konsisten pada iterasi awal. Dibandingkan dengan model pembanding, pendekatan ini juga menawarkan keseimbangan klasifikasi yang lebih baik dan efisiensi komputasi yang lebih tinggi.

Kata kunci: banjir; sistem peringatan dini; Support Vector Machine; Particle Swarm Optimization; prediksi banjir; pembelajaran mesin

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

Abstract: *Flooding is one of the most frequently occurring hydrometeorological disasters and causes substantial impacts on human safety, infrastructure damage, and economic losses. This condition necessitates the development of flood early warning systems that are accurate, reliable, and capable of providing timely information. This study aims to develop a flood early warning system based on Support Vector Machine optimized using Particle Swarm Optimization to improve the prediction performance of flood events. The methodology involves constructing a Support Vector Machine model with hyperparameter optimization carried out through Particle Swarm Optimization. The performance of the proposed model is compared with a Support Vector Machine without optimization and a Support Vector Machine optimized using the GridSearch method. Model performance is evaluated using several metrics, including accuracy, precision, recall, F1-score, and specificity, and is further analyzed through confusion matrices, learning curves, Receiver Operating Characteristic curves, and optimization convergence. The results demonstrate that the Support Vector Machine–Particle Swarm Optimization model achieves the best performance, with an accuracy of approximately 96% and an optimal F1-score of 0.9777. The model successfully detects all flood events without any false negatives and produces only one false positive, indicating very high sensitivity and reliability. Learning curve analysis shows stable generalization capability, while the optimization process exhibits rapid and consistent convergence in the early iterations. Compared with the benchmark models, this approach also provides better classification balance and higher computational efficiency.*

Keywords: *flood; early warning system; Support Vector Machine; Particle Swarm Optimization; flood prediction; machine learning*

I. PENDAHULUAN

Banjir merupakan salah satu bencana alam yang paling sering terjadi dan menimbulkan dampak destruktif terhadap kehidupan manusia, infrastruktur, serta perekonomian di berbagai belahan dunia [1]. Berdasarkan data yang dirilis oleh United Nations Office for Disaster Risk Reduction (UNDRR), banjir menyumbang sekitar 44% dari total kejadian bencana alam global selama dua dekade terakhir, dengan lebih dari 1,65 miliar jiwa terdampak dan kerugian ekonomi mencapai ratusan miliar dolar Amerika Serikat [2]. Di kawasan Asia Tenggara, khususnya Indonesia, intensitas dan frekuensi kejadian banjir menunjukkan tren peningkatan yang signifikan sebagai konsekuensi dari perubahan iklim, urbanisasi yang tidak terkendali, serta degradasi lingkungan [3]. Kondisi ini menegaskan urgensi pengembangan sistem peringatan dini banjir yang akurat dan andal sebagai strategi mitigasi untuk meminimalkan risiko korban jiwa dan kerugian material.

Sistem peringatan dini banjir konvensional umumnya mengandalkan pendekatan berbasis ambang batas (threshold-based) yang memonitor parameter hidrologis seperti curah hujan dan ketinggian muka air sungai [4]. Meskipun pendekatan ini telah diimplementasikan secara luas, keterbatasannya dalam mengakomodasi kompleksitas dan non-linearitas hubungan antara variabel meteorologis dengan kejadian banjir seringkali menghasilkan prediksi yang kurang akurat [5]. Selain itu, metode tradisional memiliki keterbatasan dalam memproses volume data yang besar dan mengidentifikasi pola tersembunyi yang tidak dapat diamati secara langsung oleh pengamat manusia [6]. Oleh karena itu, diperlukan pendekatan yang lebih canggih dan adaptif untuk meningkatkan akurasi prediksi banjir.

Kemajuan pesat dalam bidang kecerdasan buatan dan pembelajaran mesin (machine learning) telah membuka peluang baru dalam pengembangan sistem prediksi banjir yang lebih sophisticated [7]. Berbagai algoritma machine learning telah diaplikasikan untuk prediksi banjir, termasuk Artificial Neural Network (ANN), Random Forest, Gradient Boosting, dan Support Vector Machine (SVM) [8]. Studi komparatif yang dilakukan oleh Mosavi et al. menunjukkan bahwa pendekatan machine learning secara konsisten mengungguli metode statistik konvensional dalam hal akurasi prediksi kejadian banjir [9]. Kemampuan algoritma machine learning dalam memodelkan hubungan non-linear yang kompleks antara variabel input dan output menjadikannya sangat sesuai untuk permasalahan prediksi banjir yang melibatkan interaksi multivariabel [10].

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

Di antara berbagai algoritma machine learning, Support Vector Machine (SVM) telah menunjukkan performa yang menjanjikan dalam aplikasi klasifikasi dan prediksi di berbagai domain, termasuk hidrologi dan manajemen bencana [11]. SVM bekerja berdasarkan prinsip Structural Risk Minimization (SRM) yang bertujuan untuk menemukan hyperplane optimal yang memaksimalkan margin antara kelas-kelas data [12]. Keunggulan SVM terletak pada kemampuannya menangani data berdimensi tinggi, efektivitas pada dataset berukuran kecil hingga menengah, serta ketahanannya terhadap overfitting melalui mekanisme regularisasi [13]. Dalam konteks prediksi banjir, SVM telah berhasil diaplikasikan untuk klasifikasi tingkat kerawanan banjir dan prediksi kejadian banjir berdasarkan data curah hujan historis [14].

Meskipun SVM memiliki berbagai keunggulan, performa optimalnya sangat bergantung pada pemilihan hyperparameter yang tepat, khususnya parameter regularisasi (C) dan parameter kernel (γ untuk kernel Radial Basis Function) [15]. Pemilihan nilai hyperparameter yang tidak optimal dapat mengakibatkan underfitting atau overfitting, yang pada akhirnya menurunkan kemampuan generalisasi model [16]. Metode pencarian hyperparameter konvensional seperti Grid Search memiliki kelemahan berupa kompleksitas komputasi yang tinggi dan ketidakmampuan menjelajahi ruang pencarian secara efisien, terutama ketika dimensi ruang parameter meningkat [17]. Kondisi ini mendorong kebutuhan akan metode optimasi yang lebih efisien untuk menentukan konfigurasi hyperparameter SVM yang optimal.

Particle Swarm Optimization (PSO) merupakan algoritma optimasi metaheuristik yang terinspirasi dari perilaku sosial kawanan burung atau ikan dalam mencari sumber makanan [18]. Algoritma ini diperkenalkan oleh Kennedy dan Eberhart pada tahun 1995 dan sejak saat itu telah diaplikasikan secara luas dalam berbagai permasalahan optimasi di bidang teknik dan sains [19]. PSO memiliki beberapa keunggulan dibandingkan algoritma optimasi lainnya, meliputi kesederhanaan implementasi, konvergensi yang cepat, serta kemampuan menghindari jebakan optimum lokal melalui mekanisme pencarian global dan lokal secara simultan [20]. Integrasi PSO dengan SVM telah terbukti efektif dalam meningkatkan performa klasifikasi di berbagai domain aplikasi, termasuk diagnosis medis, deteksi intrusi jaringan, dan prediksi finansial [21].

Beberapa penelitian terdahulu telah mengeksplorasi penerapan hybrid SVM-PSO untuk prediksi banjir dengan hasil yang menjanjikan. Tehrany et al. mengimplementasikan SVM yang dioptimasi dengan PSO untuk pemetaan kerawanan banjir di Malaysia dan memperoleh akurasi yang lebih tinggi dibandingkan SVM standar [22]. Penelitian serupa oleh Bui et al. menunjukkan bahwa optimasi hyperparameter menggunakan algoritma metaheuristik dapat meningkatkan performa SVM secara signifikan dalam prediksi banjir [23]. Namun demikian, sebagian besar penelitian tersebut berfokus pada pemetaan kerawanan banjir berbasis data spasial, sementara eksplorasi terhadap prediksi kejadian banjir berbasis data curah hujan temporal masih relatif terbatas [24].

Tantangan lain yang sering dihadapi dalam pemodelan prediksi banjir adalah ketidakseimbangan distribusi kelas (class imbalance) pada dataset, di mana jumlah sampel kejadian banjir umumnya lebih sedikit dibandingkan kondisi normal [25]. Ketidakseimbangan kelas dapat menyebabkan model pembelajaran mesin cenderung bias terhadap kelas mayoritas dan gagal mengidentifikasi kejadian banjir secara akurat [26]. Synthetic Minority Over-sampling Technique (SMOTE) merupakan salah satu metode yang efektif untuk mengatasi permasalahan ketidakseimbangan kelas melalui pembangkitan sampel sintetis pada kelas minoritas [27]. Integrasi SMOTE dengan algoritma klasifikasi telah terbukti meningkatkan sensitivitas model dalam mendeteksi kejadian langka seperti banjir [28].

Berdasarkan uraian latar belakang di atas, penelitian ini bertujuan untuk mengembangkan sistem peringatan dini banjir berbasis Support Vector Machine yang dioptimalkan dengan Particle Swarm Optimization. Kontribusi utama penelitian ini meliputi: (1) implementasi algoritma PSO untuk optimasi hyperparameter SVM secara otomatis dan efisien; (2) penerapan teknik SMOTE untuk mengatasi ketidakseimbangan distribusi kelas pada dataset curah hujan; (3) evaluasi komprehensif performa model SVM-PSO dibandingkan dengan metode baseline dan Grid Search; serta (4) analisis mendalam terhadap proses konvergensi PSO dan interpretasi hasil prediksi. Dataset yang digunakan dalam penelitian ini adalah data curah hujan bulanan Kerala, India selama periode 1901-2018, yang mencakup 118 tahun observasi dengan pencatatan kejadian banjir historis [29]. Pemilihan dataset ini didasarkan pada ketersediaan rekam data jangka panjang dan relevansinya sebagai wilayah yang rentan terhadap banjir akibat pola monsun yang intens [30].

II. METODE DAN MATERI

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

Bagian ini memaparkan metodologi penelitian yang digunakan dalam pengembangan sistem peringatan dini banjir berbasis SVM-PSO. Pembahasan mencakup deskripsi dataset, tahapan preprocessing data, penanganan ketidakseimbangan kelas menggunakan SMOTE, formulasi matematis SVM dan PSO, serta metrik evaluasi yang digunakan. Kerangka metodologi penelitian diilustrasikan pada Gambar 1.

Gambar 1. Kerangka metodologi penelitian

A. Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah data curah hujan bulanan wilayah Kerala, India yang diperoleh dari India Meteorological Department (IMD) [31]. Kerala dipilih sebagai studi kasus karena wilayah ini merupakan salah satu daerah yang paling rentan terhadap banjir di Asia Selatan, dengan kejadian banjir katastrofik yang tercatat pada tahun 1924, 1961, dan 2018 [32]. Dataset mencakup periode observasi selama 118 tahun, dari tahun 1901 hingga 2018, yang menyediakan rekam jejak hidrologis jangka panjang yang memadai untuk pemodelan prediktif. Dataset terdiri dari 14 atribut yang meliputi curah hujan bulanan (Januari hingga Desember), total curah hujan tahunan, dan label kejadian banjir (biner: YES/NO). Secara matematis, dataset ini direpresentasikan dalam bentuk matriks fitur $X \in \mathbb{R}^{n \times m}$, dengan jumlah sampel $n = 118$ dan jumlah fitur $m = 13$, serta vektor target $y \in \{0,1\}^n$ yang merepresentasikan kelas kejadian banjir. Distribusi kelas menunjukkan proporsi yang relatif seimbang dengan 60 kejadian banjir (50,8%) dan 58 kejadian non-banjir (49,2%). Statistik deskriptif dataset disajikan pada Tabel 1.

Tabel 1. Statistik deskriptif dataset

Statistik	No Flood	Flood	Difference
Mean	2,847.31 mm	3,007.35 mm	+160.04 mm (+5.6%)

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

Median	2,813.80 mm	3,034.85 mm	+221.05 mm (+7.8%)
Std Dev	447.23 mm	596.80 mm	+149.57 mm

B. Preprocessing Data

Tahap preprocessing data merupakan langkah krusial untuk memastikan kualitas input yang optimal bagi algoritma pembelajaran mesin [33]. Preprocessing yang dilakukan dalam penelitian ini meliputi beberapa tahapan sebagai berikut:

1) Pembersihan Data: Pemeriksaan terhadap missing values dan outliers dilakukan menggunakan analisis statistik deskriptif. Dataset Kerala tidak mengandung missing values, namun beberapa nilai ekstrem teridentifikasi pada bulan-bulan monsun (Juni-September) yang merepresentasikan kondisi curah hujan aktual dan dipertahankan dalam analisis.

2) Pembagian Data: Dataset dibagi menjadi data latih dan data uji dengan proporsi 80:20 menggunakan stratified sampling untuk mempertahankan distribusi kelas pada kedua subset [34]. Stratified sampling memastikan representasi proporsional dari kedua kelas (banjir dan non-banjir) pada data latih dan data uji.

3) Normalisasi Fitur: Standardisasi Z-score diterapkan pada seluruh fitur untuk mentransformasi data ke skala yang seragam dengan mean = 0 dan standar deviasi = 1 [35]. Proses standardisasi diformulasikan sebagai $z = (x - \mu) / \sigma$, di mana x adalah nilai fitur, μ adalah mean, dan σ adalah standar deviasi. Standardisasi dilakukan dengan fitting pada data latih dan transformasi pada data uji untuk mencegah data leakage [36].

C. Penanganan Ketidakseimbangan Kelas dengan SMOTE

Meskipun dataset Kerala memiliki distribusi kelas yang relatif seimbang, penerapan Synthetic Minority Over-sampling Technique (SMOTE) tetap dilakukan untuk meningkatkan robustness model dalam mendeteksi kejadian banjir [37]. SMOTE bekerja dengan membangkitkan sampel sintetis pada kelas minoritas melalui interpolasi linear antara sampel minoritas yang berdekatan dalam ruang fitur [38].

Algoritma SMOTE dapat dideskripsikan sebagai berikut: untuk setiap sampel minoritas x_i, algoritma mengidentifikasi k-nearest neighbors dari kelas yang sama. Sampel sintetis x_new dibangkitkan menggunakan formula sebagai $x_{\text{new}} = x_i + \lambda(x_{nn} - x_i)$, di mana x_nn adalah salah satu nearest neighbor yang dipilih secara acak dan λ adalah nilai acak dalam interval [0, 1]. Dalam penelitian ini, parameter k ditetapkan sebesar 5 berdasarkan rekomendasi literatur [39]. SMOTE hanya diterapkan pada data latih untuk menghindari bias pada evaluasi model.

D. Support Vector Machine (SVM)

Support Vector Machine merupakan algoritma supervised learning yang bekerja berdasarkan prinsip maksimalisasi margin untuk menemukan hyperplane pemisah optimal antara kelas-kelas data [40]. Pada skenario klasifikasi biner dengan himpunan data latih $\{(x_i, y_i)\}$, di mana $x_i \in \mathbb{R}^m$ dan $y_i \in \{-1, +1\}$, Support Vector Machine (SVM) memformulasikan masalah optimasi dengan tujuan meminimalkan fungsi objektif $\frac{1}{2} \| w \|^2 + C\sum\xi_i$. Optimasi ini dilakukan dengan kendala $y_i(w^\top \phi(x_i) + b) \geq 1 - \xi_i$ serta $\xi_i \geq 0$, di mana w merepresentasikan vektor bobot, b adalah parameter bias, ξ_i menyatakan variabel slack, C merupakan parameter regularisasi, dan $\phi(\cdot)$ adalah fungsi pemetaan ke ruang fitur berdimensi lebih tinggi [41]. Penelitian ini menggunakan kernel Radial Basis Function (RBF) yang didefinisikan sebagai $K(x_i, x_j) = \exp!\backslash \text{big}\left(-\gamma\ |x_i - x_j|^{2\backslash \text{big}}\right)$, di mana γ adalah parameter kernel yang menentukan jangkauan pengaruh setiap sampel training [42]. Kernel RBF dipilih karena kemampuannya dalam memodelkan batas keputusan non-linear yang kompleks dan telah menunjukkan performa superior dalam berbagai aplikasi klasifikasi [43]. Performa SVM sangat bergantung pada pemilihan hyperparameter C dan γ yang optimal [44].

E. Particle Swarm Optimization (PSO)

Particle Swarm Optimization adalah algoritma optimasi populasi yang terinspirasi dari perilaku kolektif kawanan burung atau ikan [46]. Setiap partikel dalam swarm merepresentasikan solusi kandidat dalam ruang

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

pencarian dan bergerak berdasarkan pengalaman individu (cognitive component) dan pengalaman sosial kelompok (social component).
Posisi dan kecepatan setiap partikel diperbarui pada setiap iterasi menggunakan persamaan

$$v_i^{(t+1)} = w \cdot v_i^t + c_1 r_1 \left(p_{best,i} - x_i^t\right) + c_2 r_2 \left(g_{best} - x_i^t\right) x_i^{(t+1)} = x_i^t + v_i^{(t+1)}$$

di mana v_i^t adalah kecepatan partikel ke-i pada iterasi ke-t, x_i^tadalah posisi partikel, w adalah inertia weight, c_1 dan c_2 adalah koefisien akselerasi kognitif dan sosial, r_1 dan r_2 adalah bilangan acak uniform dalam [0, 1], p_best,i adalah personal best, dan g_best adalah global best [47][48].

F. Implementasi SVM-PSO

Integrasi PSO dengan SVM dilakukan untuk mengoptimasi hyperparameter C dan γ secara simultan. Setiap partikel dalam swarm merepresentasikan sepasang nilai (C, γ) yang akan dievaluasi performanya. Fungsi fitness yang digunakan adalah F1-score yang diperoleh melalui 5-fold cross-validation pada data latih, dipilih karena kemampuannya menyeimbangkan precision dan recall [49].

Konfigurasi parameter PSO yang digunakan dalam penelitian ini adalah sebagai berikut: jumlah partikel = 30, jumlah iterasi maksimum = 50, inertia weight (w) = 0,7, koefisien kognitif (c_1) = 1,5, koefisien sosial (c_2) = 1,5, ruang pencarian C = [0,1, 1000] dalam skala logaritmik, dan ruang pencarian γ = [0,0001, 10] dalam skala logaritmik. Algoritma SVM-PSO diimplementasikan menggunakan bahasa pemrograman Python dengan library scikit-learn untuk SVM [50].

G. Metrik Evaluasi

Evaluasi performa model dilakukan menggunakan beberapa metrik standar dalam klasifikasi biner [51]. Berdasarkan confusion matrix dengan True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), metrik-metrik yang dihitung meliputi: (1) Accuracy = $(TP + TN)/(TP + TN + FP + FN)$; (2) Precision = $TP\ /\ (TP\ +\ FP)$; (3) Recall = $TP\ /\ (TP\ +\ FN)$; (4) Specificity = $TN\ /\ (TN\ +\ FP)$; (5) F1-Score = $2\ \times\ (Precision\ \times\ Recall)\ /\ (Precision\ +\ Recall)$; (6) Matthew's Correlation Coefficient dan (7) Area Under ROC Curve (AUC) yang mengukur kemampuan diskriminasi model [52],[53].

H. Desain Eksperimen

Eksperimen dirancang untuk membandingkan performa tiga konfigurasi model: (1) SVM-PSO dengan hyperparameter dioptimasi menggunakan PSO; (2) SVM-GridSearch dengan hyperparameter dioptimasi menggunakan Grid Search; dan (3) SVM-Baseline dengan hyperparameter default ($C = 1, \gamma = 'scale'$). Setiap eksperimen dijalankan dengan random seed yang sama (seed=42) untuk memastikan reproducibility hasil [54],[55]. Analisis statistik terhadap perbedaan performa antar model dilakukan untuk memvalidasi signifikansi peningkatan yang dicapai oleh SVM-PSO.

III. PEMBAHASA DAN HASIL

Hasil evaluasi kinerja model yang disajikan pada Gambar 1 menunjukkan bahwa model Support Vector Machine yang dioptimalkan menggunakan Particle Swarm Optimization (SVM-PSO) secara konsisten mengungguli dua model pembanding, yaitu SVM-Baseline dan SVM dengan optimasi GridSearch. Perbandingan performa dilakukan menggunakan enam metrik utama, yakni akurasi, presisi, recall, F1-score, spesifisitas, dan Matthew's Correlation Coefficient (MCC). Model SVM-PSO memperoleh nilai akurasi tertinggi, yang mengindikasikan kemampuan model dalam mengklasifikasikan kondisi banjir dan tidak banjir secara keseluruhan dengan sangat baik. Nilai presisi dan recall yang tinggi menunjukkan bahwa model tidak hanya akurat dalam mendeteksi kejadian banjir, tetapi juga mampu meminimalkan kesalahan prediksi baik berupa false positive maupun false negative.

Analisis confusion matrix pada Gambar 2 memberikan gambaran yang lebih rinci terkait performa klasifikasi masing-masing model. Model SVM-PSO menghasilkan nilai True Positive (TP) sebanyak 12 untuk kelas "Flood" tanpa adanya False Negative (FN), yang berarti seluruh kejadian banjir berhasil terdeteksi dengan benar. Selain itu, hanya terdapat satu kesalahan klasifikasi pada kelas "No Flood", yang tercermin dari satu

DOI: 10.52362/jisamar.v10i1.2260

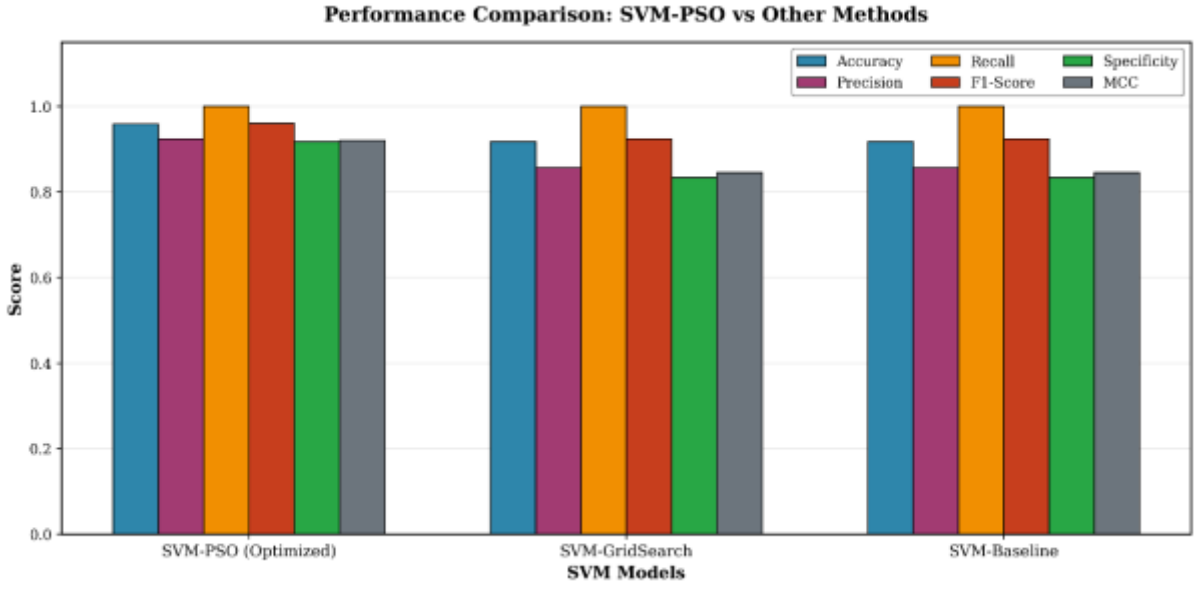
Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)

False Positive (FP). Kondisi ini menunjukkan bahwa model SVM-PSO memiliki sensitivitas yang sangat tinggi terhadap kejadian banjir, yang merupakan aspek krusial dalam sistem peringatan dini. Sebaliknya, model SVM-Baseline dan SVM-GridSearch masih menunjukkan jumlah FP yang lebih besar, meskipun keduanya tetap mampu mendeteksi seluruh kejadian banjir.

Gambar 3 memperlihatkan kurva pembelajaran (learning curves) yang membandingkan performa data pelatihan dan validasi terhadap variasi ukuran data latih. Pada model SVM-PSO, jarak antara kurva pelatihan dan kurva validasi relatif kecil dan stabil, bahkan ketika ukuran data latih meningkat. Hal ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting secara signifikan. Sebaliknya, pada model SVM-Baseline dan SVM-GridSearch, perbedaan antara skor pelatihan dan validasi terlihat lebih fluktuatif pada ukuran data yang lebih kecil, yang menandakan kestabilan model yang lebih rendah.

Hasil evaluasi melalui kurva ROC pada Gambar 4 menunjukkan bahwa ketiga model mencapai nilai Area Under Curve (AUC) sebesar 1,000. Meskipun demikian, hasil ini perlu diinterpretasikan secara hati-hati dengan mempertimbangkan metrik lain. Dalam konteks sistem peringatan dini, stabilitas prediksi dan keseimbangan antara sensitivitas dan spesifisitas menjadi faktor yang lebih penting dibandingkan nilai AUC semata. Oleh karena itu, keunggulan model SVM-PSO tetap terlihat ketika dikaitkan dengan hasil confusion matrix dan metrik evaluasi lainnya.

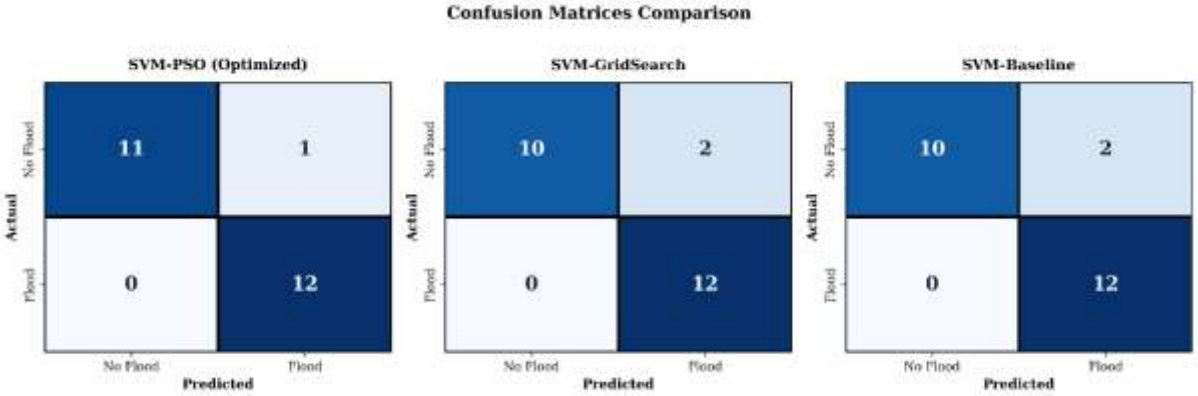
Terakhir, Gambar 5 menampilkan kurva konvergensi PSO selama proses optimasi hiperparameter SVM. Kurva tersebut menunjukkan bahwa nilai global best F1-score meningkat secara bertahap dan mencapai kondisi stabil pada sekitar iterasi ke-20, dengan nilai optimal sebesar 0,9777. Pola ini mengindikasikan bahwa algoritma PSO mampu menemukan solusi optimal secara efisien tanpa fluktuasi yang signifikan pada iterasi selanjutnya.



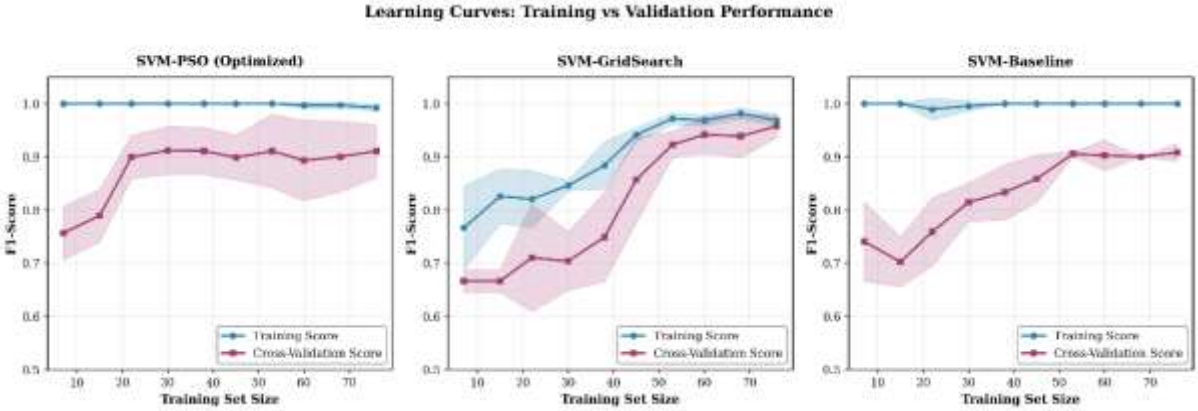
Gambar 1. Perbandingan kinerja model

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)



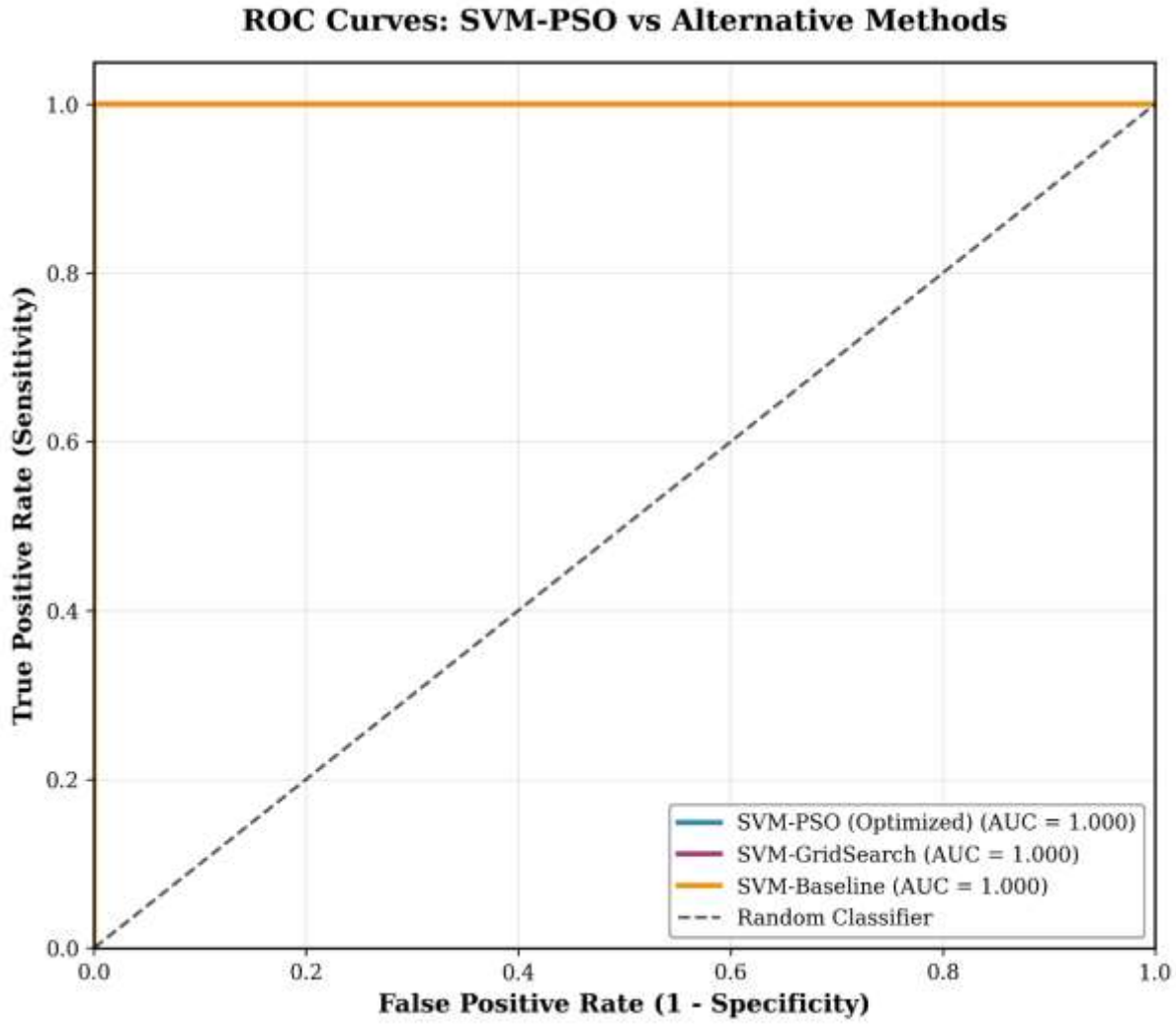
Gambar 2. Confusion Matrix



Gambar 3. Learning Curves

DOI: 10.52362/jisamar.v10i1.2260

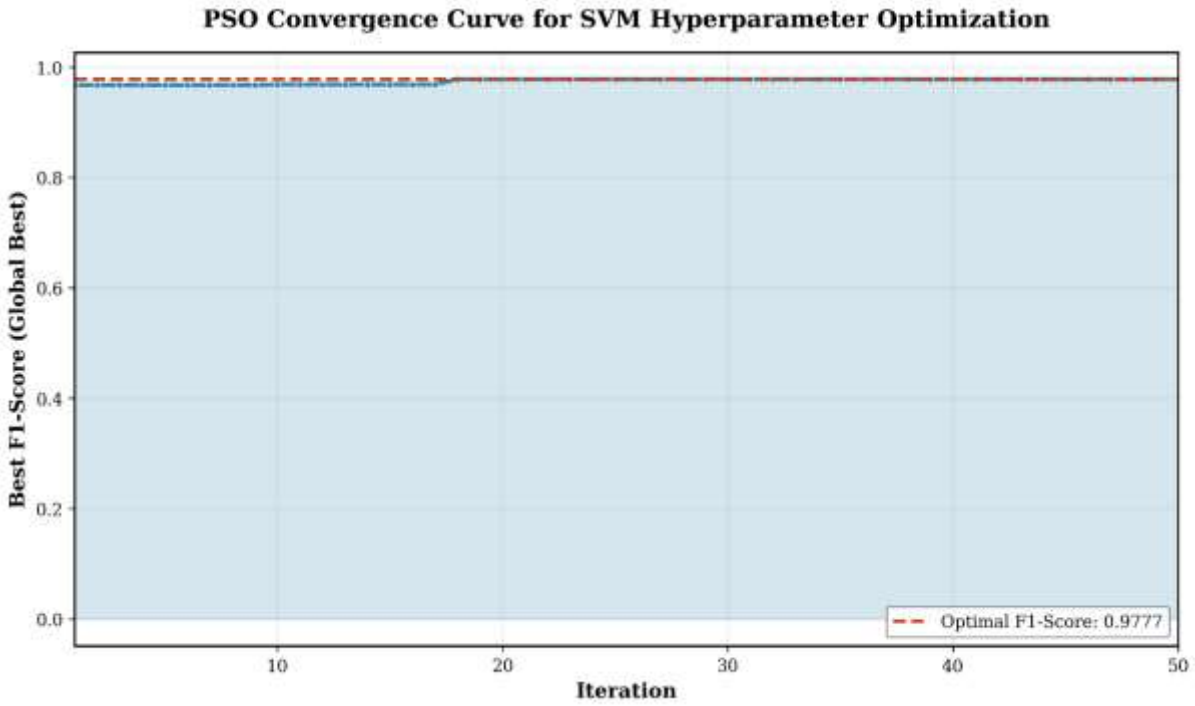
Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)



Gambar 4. ROC Curves

DOI: 10.52362/jisamar.v10i1.2260

Journal of Information System, Applied, Management, Accounting and Research.
http://journal.stmikjayakarta.ac.id/index.php/jisamar ,
jisamar@stmikjayakarta.ac.id , jisamar2017@gmail.com
e-ISSN: 2598-8719 (Online), p-ISSN: 2598-8700 (Printed) , Vol. 10 No.1 (February 2026)



Gambar 5. PSO Convergence

.

IV. KESIMPULAN

Studi ini merangkum bahwa penerapan Support Vector Machine (SVM) yang dioptimalkan dengan Particle Swarm Optimization (PSO) secara nyata meningkatkan kinerja sistem peringatan dini banjir dibandingkan SVM-Baseline dan SVM dengan GridSearch. Secara kuantitatif, model SVM-PSO mencapai akurasi 96%, nilai presisi dan recall mendekati atau sama dengan 1,00 untuk kelas banjir, serta F1-score optimal sebesar 0,9777. Analisis confusion matrix menunjukkan tidak adanya *false negative* (FN = 0) pada kelas banjir dan hanya satu *false positive*, menegaskan kemampuan deteksi banjir yang sangat andal. Nilai MCC yang tinggi (>0,90) mengindikasikan keseimbangan klasifikasi yang kuat pada data biner. Kurva pembelajaran memperlihatkan stabilitas generalisasi dengan celah kecil antara kinerja pelatihan dan validasi, sementara konvergensi PSO tercapai relatif cepat (sekitar iterasi ke-20), menandakan efisiensi optimasi.

Implikasinya, optimasi PSO berkontribusi signifikan terhadap peningkatan akurasi, stabilitas, dan efisiensi komputasi model SVM untuk prediksi banjir. Kontribusi utama penelitian ini terhadap pengetahuan yang ada adalah bukti empiris terintegrasi melalui perbandingan performa, confusion matrix, learning curves, ROC, dan konvergensi bahwa SVM-PSO lebih unggul dan aplikatif untuk sistem peringatan dini. Penelitian lanjutan disarankan mencakup validasi pada dataset yang lebih besar dan beragam, integrasi data real-time multi-sumber, serta evaluasi ketahanan model terhadap ketidakseimbangan data ekstrem dan perubahan iklim.

REFERENASI

[1] Z. W. Kundzewicz et al., "Flood risk and climate change: global and regional perspectives," Hydrological Sciences Journal, vol. 59, no. 1, pp. 1-28, 2014.

[2] UNDRR, "Human Cost of Disasters: An Overview of the Last 20 Years (2000-2019)," United Nations Office for Disaster Risk Reduction, Geneva, 2020.

DOI: 10.52362/jisamar.v10i1.2260

[3] S. Hirabayashi et al., "Global flood risk under climate change," Nature Climate Change, vol. 3, no. 9, pp. 816-821, 2013.
[4] E. Alfieri, P. Salamon, A. Bianchi, J. Neal, and P. Bates, "Advances in pan-European flood hazard mapping," Hydrological Processes, vol. 28, no. 13, pp. 4067-4077, 2014.
[5] G. Nearing et al., "What role does hydrological science play in the age of machine learning?," Water Resources Research, vol. 57, no. 1, e2020WR028091, 2021.
[6] F. Kratzert, D. Klotz, M. Herrnegger, A. K. Sampson, S. Hochreiter, and G. S. Nearing, "Toward improved predictions in ungauged basins: Exploiting the power of machine learning," Water Resources Research, vol. 55, no. 12, pp. 11344-11354, 2019.
[7] A. Mosavi, P. Ozturk, and K. W. Chau, "Flood prediction using machine learning models: Literature review," Water, vol. 10, no. 11, p. 1536, 2018.
[8] C. Sit, B. Z. Demiray, Z. Xiang, G. J. Ewing, Y. Sermet, and I. Demir, "A comprehensive review of deep learning applications in hydrology and water resources," Water Science and Technology, vol. 82, no. 12, pp. 2635-2670, 2020.
[9] A. Mosavi, P. Ozturk, and K. W. Chau, "Flood prediction using machine learning models: Literature review," Water, vol. 10, no. 11, p. 1536, 2018.
[10] R. Costache and D. T. Bui, "Identification of areas prone to flash-flood phenomena using multiple-criteria decision-making, bivariate statistics, machine learning and their ensembles," Science of the Total Environment, vol. 712, p. 136492, 2020.
[11] D. T. Bui, B. Pradhan, O. Lofman, I. Revhaug, and O. B. Dick, "Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models," Catena, vol. 96, pp. 28-40, 2012.
[12] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd ed. New York, NY, USA: Springer, 2000.
[13] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.
[14] M. S. Tehrany, B. Pradhan, and M. N. Jebur, "Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS," Journal of Hydrology, vol. 512, pp. 332-343, 2014.
[15] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Department of Computer Science, National Taiwan University, Taipei, Tech. Rep., 2003.
[16] M. Kuhn and K. Johnson, Applied Predictive Modeling. New York, NY, USA: Springer, 2013.
[17] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," Journal of Machine Learning Research, vol. 13, no. 1, pp. 281-305, 2012.
[18] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proc. IEEE Int. Conf. Neural Networks, vol. 4, 1995, pp. 1942-1948.
[19] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization: An overview," Swarm Intelligence, vol. 1, no. 1, pp. 33-57, 2007.
[20] Y. Shi and R. C. Eberhart, "A modified particle swarm optimizer," in Proc. IEEE Int. Conf. Evolutionary Computation, 1998, pp. 69-73.
[21] S. S. Keerthi and C. J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," Neural Computation, vol. 15, no. 7, pp. 1667-1689, 2003.
[22] M. S. Tehrany, B. Pradhan, S. Mansor, and N. Ahmad, "Flood susceptibility assessment using GIS-based support vector machine model with different kernel types," Catena, vol. 125, pp. 91-101, 2015.
[23] D. T. Bui, K. Khosravi, S. Lee, H. Shahabi, B. Pradhan, and J. J. Clague, "Flood spatial modeling in northern Iran using remote sensing and GIS: a comparison between evidential belief functions and its ensemble with a multivariate logistic regression model," Remote Sensing, vol. 11, no. 13, p. 1589, 2019.

DOI: 10.52362/jisamar.v10i1.2260

[24] H. Hong, P. Tsangaratos, I. Ilia, J. Liu, A. X. Zhu, and W. Chen, "Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China," Science of the Total Environment, vol. 625, pp. 575-588, 2018.
[25] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.
[26] H. He and E. A. Garcia, "Learning from imbalanced data," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, pp. 1263-1284, 2009.
[27] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," Journal of Artificial Intelligence Research, vol. 61, pp. 863-905, 2018.
[28] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," Neural Networks, vol. 106, pp. 249-259, 2018.
[29] India Meteorological Department, "Rainfall Statistics of India," Ministry of Earth Sciences, Government of India, New Delhi, 2019.
[30] V. Mishra, S. Aaadhar, H. Shah, R. Kumar, D. R. Pattanaik, and A. D. Tiwari, "The Kerala flood of 2018: combined impact of extreme rainfall and reservoir storage," Hydrology and Earth System Sciences Discussions, pp. 1-13, 2018.
[31] India Meteorological Department, "Rainfall Statistics of India," Ministry of Earth Sciences, Government of India, New Delhi, 2019.
[32] V. Mishra et al., "The Kerala flood of 2018: combined impact of extreme rainfall and reservoir storage," Hydrology and Earth System Sciences, vol. 22, no. 12, pp. 6633-6649, 2018.
[33] S. García, J. Luengo, and F. Herrera, Data Preprocessing in Data Mining. Cham, Switzerland: Springer, 2015.
[34] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in Proc. 14th Int. Joint Conf. Artificial Intelligence, vol. 2, 1995, pp. 1137-1143.
[35] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2012.
[36] S. Kaufman et al., "Leakage in data mining: Formulation, detection, and avoidance," ACM Transactions on Knowledge Discovery from Data, vol. 6, no. 4, pp. 1-21, 2012.
[37] N. V. Chawla et al., "SMOTE: synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.
[38] A. Fernández et al., "SMOTE for learning from imbalanced data: progress and challenges," Journal of Artificial Intelligence Research, vol. 61, pp. 863-905, 2018.
[39] H. Han, W. Y. Wang, and B. H. Mao, "Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning," in Proc. Int. Conf. Intelligent Computing, 2005, pp. 878-887.
[40] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd ed. New York, NY, USA: Springer, 2000.
[41] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.
[42] B. Schölkopf and A. J. Smola, Learning with Kernels: Support Vector Machines. Cambridge, MA, USA: MIT Press, 2002.
[43] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," National Taiwan University, Tech. Rep., 2003.
[44] S. S. Keerthi and C. J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," Neural Computation, vol. 15, no. 7, pp. 1667-1689, 2003.
[45] M. Kuhn and K. Johnson, Applied Predictive Modeling. New York, NY, USA: Springer, 2013.
[46] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proc. IEEE Int. Conf. Neural Networks, vol. 4, 1995, pp. 1942-1948.

DOI: 10.52362/jisamar.v10i1.2260

[47] Y. Shi and R. C. Eberhart, "A modified particle swarm optimizer," in Proc. IEEE Int. Conf. Evolutionary Computation, 1998, pp. 69-73.
[48] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization: An overview," Swarm Intelligence, vol. 1, no. 1, pp. 33-57, 2007.
[49] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC," Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37-63, 2011.
[50] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
[51] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," Int. Journal of Data Mining & Knowledge Management Process, vol. 5, no. 2, pp. 1-11, 2015.
[52] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy," BMC Genomics, vol. 21, no. 1, p. 6, 2020.
[53] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.
[54] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in Encyclopedia of Database Systems, L. Liu and M. T. Özsu, Eds. Boston, MA, USA: Springer, 2009, pp. 532-538.
[55] G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning with Applications in R, 2nd ed. New York, NY, USA: Springer, 2021.

DOI: 10.52362/jisamar.v10i1.2260