

Implementasi Algoritma *Support Vector Machine* (SVM) untuk Deteksi Banjir

Agnes Frenica¹, Lindawati², Sopian Soim³
^{1,2,3} Politeknik Negeri Sriwijaya, Palembang, Indonesia

Email: afrenica.agnes518@gmail.com¹, lindawati@polsri.ac.id², sopiansoim@gmail.com³

Abstract – Flooding is the most common disaster in Indonesia in 2022. Based on a report from the National Disaster Management Agency (BNPB), the number of floods equalled 43.1% of the total national disaster events of around 1,524. One of the areas that was repeatedly affected by flooding from 2022 to 2023 was South Sumatra 99 times. To overcome this problem, machine learning technology can be used as a detection and early warning tool against flooding, one of which is the SVM algorithm. In this study, the performance of various SVM kernels was assessed, and the kernel with the best performance was selected for use in the flood detection system. This research categorizes the flood status with three classification labels: "safe," "alert," and "danger." Various SVM algorithms such as linear, polynomial, RBF, and sigmoid are implemented using public datasets to build a flood status prediction model. Then it will be tested with a flood detection device consisting of an Arduino Uno, NodeMCU, and various sensors such as ultrasonic, water discharge, rainfall, and temperature and humidity sensors. Evaluation measures accuracy, precision, recall, F1-Score, ROC curve, and Cross-Validation. The evaluation results showed that the SVM model with a polynomial kernel was selected as the optimal detection model compared to other kernels. This model achieved a training and testing accuracy of 1.0, a training time of only 0.0012 seconds, a prediction time of 0.0002 seconds, and precision, recall, and F1-score of 1.0. In addition, cross-validation also reached 1.0 in classifying flood data. Testing on the tool used 131 test data with an accuracy of 1.0. Classification results and sensor data are presented through an Android application, making flood monitoring easier.

Keywords - *Support Vector Machine*, Flood, Classification, Android

Intisari – Banjir merupakan bencana yang paling banyak terjadi pada tahun 2022 di Indonesia. Berdasarkan laporan Badan Nasional Penanggulangan Bencana (BNPB), jumlah bencana banjir setara 43,1% dari total kejadian bencana nasional sekitar 1.524 kejadian. Salah satu wilayah yang secara berulang terkena dampak banjir selama periode 2022 hingga 2023 adalah Sumatera Selatan sebanyak 99 kali. Dalam upaya penanggulangan masalah ini, teknologi machine learning dapat digunakan sebagai alat pendeteksi dan peringatan dini terhadap banjir, salah satunya Algoritma SVM. Dalam penelitian ini, kinerja berbagai kernel SVM dinilai, dan kernel dengan kinerja terbaik dipilih untuk digunakan dalam sistem deteksi banjir. Penelitian ini mengkategorikan status banjir dengan tiga label klasifikasi: "aman," "waspada," dan "bahaya". Dalam upaya membangun model prediksi status banjir, berbagai algoritma SVM seperti linear, polynomial, RBF, dan sigmoid diimplementasikan menggunakan dataset publik. Lalu akan diuji dengan alat deteksi banjir yang terdiri dari dari arduino uno, nodemcu, dan berbagai sensor seperti sensor ultrasonik, sensor debit air, sensor curah hujan, sensor suhu dan kelembaban. Evaluasi dilakukan melalui pengukuran akurasi, presisi, recall, F1-Score, kurva ROC, dan Cross-Validation. Hasil evaluasi menunjukkan bahwa model SVM dengan kernel polynomial terpilih sebagai model deteksi yang optimal dibandingkan dengan kernel lainnya. Model ini mencapai akurasi pelatihan dan pengujian sebesar 1.0, waktu pelatihan hanya 0.0012 detik, waktu prediksi 0.0002 detik, serta presisi, recall, dan F1-score sebesar 1.0. Selain itu, cross-validation juga mencapai 1.0 dalam mengklasifikasikan data banjir. Pengujian pada alat digunakan 131 data uji dengan hasil akurasi sebesar 1.0. Hasil klasifikasi dan data sensor disajikan melalui aplikasi Android, memudahkan dalam pemantauan banjir.

Kata Kunci – *Support Vector Machine*, Banjir, Klasifikasi, Android

I. PENDAHULUAN

Di Indonesia, banjir merupakan salah satu bencana alam yang sering terjadi [1], dengan dampak merugikan bagi masyarakat dan lingkungan. Badan Nasional Penanggulangan Bencana (BNPB) melaporkan bahwa banjir merupakan bencana yang paling sering terjadi dan melanda banyak wilayah sepanjang tahun 2022, dengan total 1.524 kejadian. Angka ini setara dengan 43,1% dari seluruh kejadian bencana nasional [2]. Beberapa faktor yang meningkatkan potensi risiko banjir meliputi sistem drainase yang tidak memadai di permukiman, kurangnya kesadaran masyarakat akan pentingnya lingkungan, minimnya upaya konservasi air, topografi dengan ketinggian tanah yang rendah, serta tingginya curah hujan. Selain itu, perkembangan kota yang tidak terkendali juga berkontribusi pada meningkatnya potensi banjir [3].

Sebagai respons terhadap permasalahan ini, berbagai upaya telah dilakukan untuk mengembangkan sistem pendeteksian dan peringatan dini banjir. Penelitian [4] menghasilkan sistem pendeteksi banjir yang menggunakan sensor potensiometer berbasis Arduino Uno. Sistem ini mengukur ketinggian air dan hanya menampilkan informasi melalui LCD, serta dilengkapi dengan lampu LED dan alarm sebagai peringatan saat banjir terdeteksi. Sistem Monitoring dan Peringatan Ketinggian Air Berbasis Web dan SMS Gateway[5] menggunakan sensor ultrasonik untuk mengukur ketinggian air dan memanfaatkan antarmuka Web dan SMS Gateway untuk memberikan informasi kepada masyarakat mengenai kondisi air. Namun, perlu diperhatikan bahwa penggunaan web hosting gratis dalam penelitian ini memiliki potensi untuk mempengaruhi respon time dan ketersediaan website, terutama dalam situasi mati listrik.

Penelitian dalam bidang Machine Learning juga memberikan sumbangan penting dalam mitigasi risiko banjir. Fokus utama dalam penelitian Machine Learning adalah bagaimana mengenali pola-pola yang kompleks secara otomatis serta mengambil keputusan cerdas berdasarkan data [6]. Ada banyak algoritma yang digunakan dalam sistem pendeteksian banjir [7] yaitu algoritma Fuzzy Logic diimplementasikan untuk memprediksi potensi bencana banjir. Dalam prosesnya, data curah hujan dan tinggi muka air diolah dengan menggunakan Algoritma Particle Swarm Optimization (PSO). Namun, penelitian ini belum melibatkan pengembangan aplikasi banjir untuk perangkat mobile, sehingga informasi mengenai prediksi banjir belum dapat diakses oleh masyarakat dengan cepat [8].

Algoritma C4.5 juga telah dimanfaatkan dalam prediksi banjir, meskipun memiliki keterbatasan dalam mengklasifikasikan data variabel hasil pengukuran untuk meningkatkan akurasi hasil [7]. Penelitian selanjutnya dilakukan pembangunan sistem deteksi banjir berbasis Arduino. Sistem ini mengintegrasikan sensor water level dan water flow dengan algoritma k-Nearest Neighbor (KNN) dan memiliki akurasi yang terbilang rendah yaitu 60% yang diakibatkan oleh kurangnya data testing yang bervariasi [9]. Pada penelitian [10][11] menggunakan Algoritma Support Vector Machine (SVM) untuk mendeteksi dan mengklasifikasi pesan bencana banjir di Twitter dengan akurasi yang didapat 96% dan 78.9%. Pengambilan data hanya dilakukan pada rentang waktu tertentu sehingga tidak mencakup semua tweet mengenai banjir yang muncul di Twitter.

Berdasarkan uraian penjelasan tersebut, peneliti tertarik untuk menggunakan Algoritma Support Vector Machine (SVM) karena pada penelitian sebelumnya [10][11], Algoritma Support Vector Machine (SVM) hanya menentukan kebenaran informasi banjir di twitter. Dalam penelitian ini, Algoritma Support Vector Machine (SVM) menjadi fokus utama. Keunggulan SVM dalam mengenali pola-pola kompleks dalam data dan mampu menerapkan pemisahan linear pada input data non-linear berdimensi tinggi [12] membuatnya menjadi pilihan yang menarik. Oleh karena itu, penelitian yang berjudul “Implementasi Algoritma Support Vector Machine (SVM) Untuk Deteksi Banjir” bertujuan untuk merancang dan menganalisis sistem deteksi banjir dengan mengimplementasikan Algoritma Support Vector Machine (SVM) yang berbasis Android, sehingga memudahkan pemantauan atau monitoring

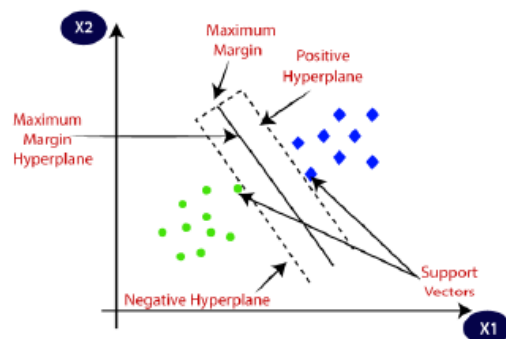
banji. Penelitian ini akan memfokuskan pada implementasi SVM untuk deteksi banjir pada simulasi buatan, tanpa mempertimbangkan variasi geografis dan iklim.

II. SIGNIFIKANSI STUDI

A. Support Vector Machine (SVM)

Pada tahun 1992, Vapnik bersama dengan rekannya Bernhard Boser dan Isabelle Guyon memperkenalkan algoritma *Support Vector Machine* (SVM) yang inovatif. SVM merupakan sebuah metode yang beroperasi dengan memanfaatkan konsep pemetaan nonlinier untuk mengubah ruang data pelatihan awal menjadi dimensi yang lebih tinggi. Dalam dimensi yang baru ini, algoritma SVM mencari hyperplane yang mampu memisahkan data dari berbagai kelas secara linier. Melalui pemetaan nonlinier yang cermat ke dimensi yang lebih tinggi ini, data dari dua kelas yang berbeda dapat selalu terpisahkan dengan jelas oleh hyperplane yang dihasilkan. Proses ini dilakukan dengan memanfaatkan konsep support vector dan margin yang ada dalam SVM [13].

Support Vector Machine (SVM) adalah salah satu algoritma populer dalam supervised learning yang digunakan dalam klasifikasi dan regresi. SVM dapat digunakan untuk memisahkan dua kelompok data dengan membuat keputusan yang lebih baik pada titik-titik data di luar set pelatihan. Keuntungan dari SVM adalah memiliki kesalahan generasi yang rendah, kompleksitas komputasi yang rendah, dan interpretasi yang mudah. Namun, penggunaan SVM juga memerlukan pemilihan kernel dan optimasi parameter yang tepat agar dapat memberikan hasil yang optimal. Dalam klasifikasi, SVM mencari hyperplane yang dapat memisahkan dua kelompok data dengan jarak maksimum, sehingga semua titik dari satu kelas berada di satu sisi hyperplane dan semua titik dari kelas lainnya berada di sisi yang berlawanan [14].



Gambar 1. Ilustrasi SVM

Pada suatu dataset, terdapat dua jenis variabel utama, yaitu variabel x_i yang merepresentasikan data, dan variabel y_i yang mengindikasikan kelas dari setiap data dalam dataset tersebut. Metode SVM (Support Vector Machine) digunakan untuk melakukan pemisahan terhadap dua kelas yang ada dalam dataset. Proses ini dilakukan dengan memanfaatkan hyperplane, suatu bidang pemisah dalam ruang berdimensi tinggi. Hyperplane ini memiliki sifat khusus: data dari kelas pertama akan ditempatkan di satu sisi hyperplane dengan nilai 1, sementara data dari kelas lainnya akan berada di sisi lain hyperplane dengan nilai -1 [15]. Hasil akurasi dari model yang dihasilkan dalam tahap transisi menggunakan SVM sangat dipengaruhi oleh jenis fungsi kernel dan parameter yang digunakan [16].

Konsep utama metode ini melibatkan pencarian ruang pemisah yang optimal di antara berbagai kelas dalam suatu dataset [17]. Ide dasarnya adalah menemukan hyperplane yang memisahkan data d -dimensional secara sempurna ke dalam dua kelasnya. Hyperplane yang dipilih SVM memiliki margin terbesar dari titik data di dua sisi. Keuntungan tambahannya

adalah bahwa algoritma hanya membutuhkan pengetahuan tentang vektor pendukung atau titik yang berada di margin [18].

Dalam tahap ini, empat jenis fungsi kernel digunakan, yaitu kernel linear, rbf, polynomial dan sigmoid. Berbagai variasi kernel yang digunakan diuraikan pada Tabel I [19].

TABEL I
TIPE KERNEL

Nama Kernel	Fungsi Kernel
Kernel Linear	$K(x_i, x) = X_i^T x$
Kernel RBF	$K(x, y) = \exp \left\{ -\frac{ x - y ^2}{2\gamma^2} \right\}$
Kernel Polynomial	$K(x_i, x) = (\gamma(x_i^T x) + r)^p$
Kernel Sigmoid	$K(x_i, x) = \tanh(\gamma(x_i^T x) + r)$

B. Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari sekumpulan data untuk mengidentifikasi atau mengklasifikasi kondisi banjir. Dataset diambil dari laman *web public* dataset yaitu Kaggle.com. Dalam penelitian ini, terdapat tiga label klasifikasi yaitu aman, waspada, dan bahaya. Kondisi banjir dapat didefinisikan berdasarkan jarak air, debit air, curah hujan, suhu, dan kelembaban. Kondisi banjir dengan status “Aman” apabila jarak air 35-100 cm, debit air 0-20 liter/detik, curah hujan 0-20 mm, suhu 30-36°C dan kelembaban 55-60%. Sedangkan untuk status “Waspada” apabila jarak air 20-34 cm, debit air 21-35 liter/detik, curah hujan 21-30 mm, suhu 25-29°C dan kelembaban 61-70%. Dan berstatus “Bahaya” apabila jarak air 0-19 cm, debit air 35-50 liter/detik, curah hujan 31-200 mm, suhu 20-24°C dan kelembaban 71-95%.

TABEL II
DEKSRIPSI DATASET

Jumlah dataset	16272
Jumlah masukan (fitur)	5
Jumlah kelas keluaran	1
Jumlah label kelas	3
Kelas	1- Aman 2- Waspada 3- Bahaya

Adapun fitur yang digunakan dan label dataset yang diuraikan pada tabel III.

TABEL III
DEKSRIPSI FITUR DAN LABEL DATASET

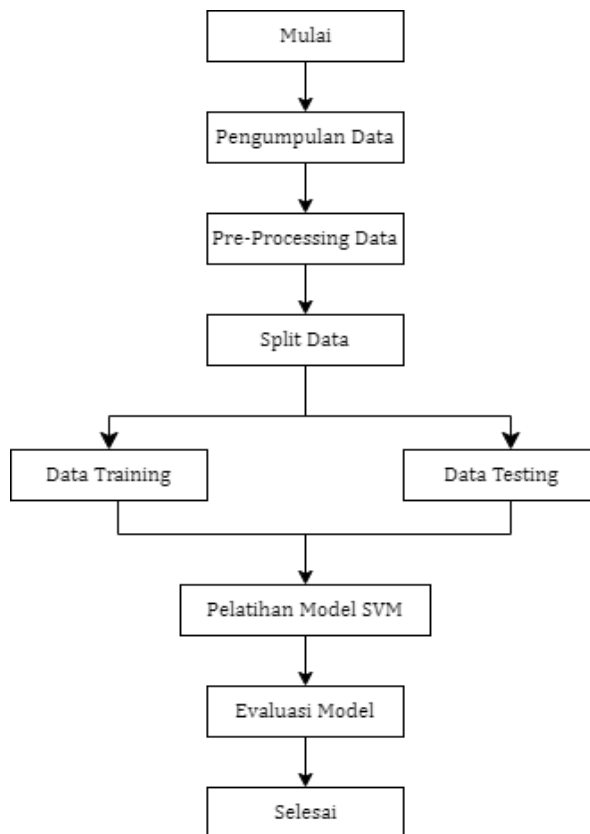
No.	Fitur	Tipe	Nilai
1	Jarak Air (cm)	Integer	[0-100]
2	Debit Air (L/detik)	Floating-Point	[0-50]
3	Curah Hujan	Floating-Point	[0-200]
4	Suhu	Floating-Point	[20-36]
5	Kelembaban	Floating-Point	[55-95]
6	Status	Integer	Aman = 0; Waspada = 1; Bahaya = 2

Berdasarkan tabel diatas, fitur yang digunakan adalah

- 1) **Jarak Air (cm):** Ini adalah fitur yang mengukur jarak air, diukur dalam sentimeter (cm). Rentang nilai dari 0 hingga 100 menunjukkan berapa dekat atau jauhnya air dari titik tertentu.
- 2) **Debit Air (L/detik):** Ini adalah fitur yang mengukur debit air, diukur dalam liter per detik (L/detik). Rentang nilai dari 0 hingga 50 menunjukkan seberapa besar aliran air dalam suatu waktu tertentu.
- 3) **Curah Hujan:** Ini adalah fitur yang mengukur curah hujan, diukur dalam satuan yang tidak disebutkan (misalnya, milimeter per waktu tertentu). Rentang nilai dari 0 hingga 200 menunjukkan seberapa banyak hujan yang turun dalam suatu waktu tertentu.
- 4) **Suhu:** Ini adalah fitur yang mengukur suhu, diukur dalam derajat Celsius (°C). Rentang nilai dari 20 hingga 36 menunjukkan kisaran suhu yang bisa terjadi.
- 5) **Kelembaban:** Ini adalah fitur yang mengukur tingkat kelembaban, diukur dalam persentase. Rentang nilai dari 55 hingga 95 menunjukkan kisaran kelembaban relatif.
- 6) **Status:** Ini adalah variabel target yang menunjukkan status keselamatan. Ini adalah variabel diskrit yang dapat memiliki nilai 0 (Aman), 1 (Waspada), atau 2 (Bahaya).

C. Metode Penelitian

Dalam penelitian ini, model prediksi status banjir menggunakan algoritma SVM dan banyak kernelnya. Empat jenis kernel yang digunakan: linear, RBF, polynomial, dan sigmoid. Penelitian ini membagi prediksi status banjir dengan tiga label kategori: “Aman”, “Waspada”, dan “Bahaya”. Adapun proses utama dari penelitian ini adalah membuat model SVM yang mampu mengklasifikasikan kondisi banjir secara akurat. Skema pelatihan model yang digunakan dalam penelitian ini adalah sebagai berikut:



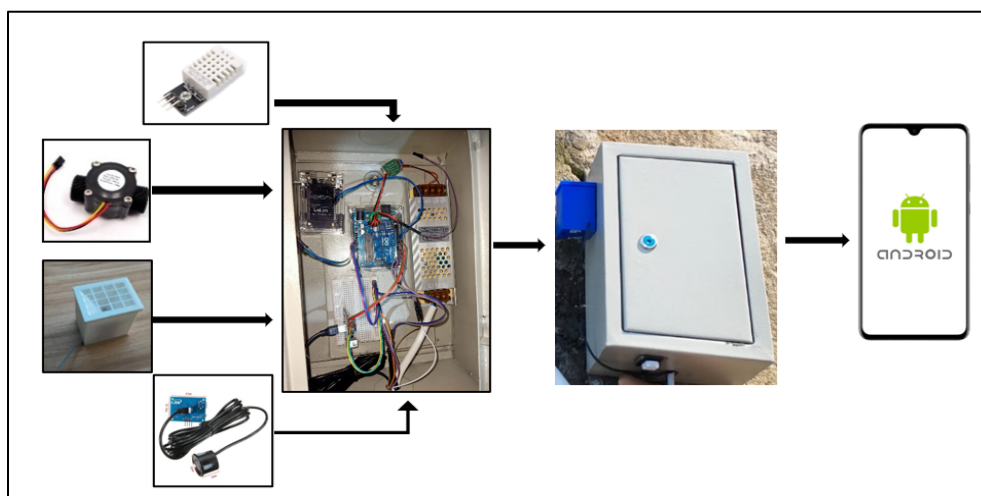
Gambar 2. Skema Pelatihan Model SVM

Proses pelatihan model *Support Vector Machine* (SVM) dimulai dengan pengumpulan dataset lalu dilakukan pre-processing data yaitu membersihkan data dari nilai yang hilang (*missing values*), memisahkan kolom gabungan, menghapus kolom yang tidak perlu, menghilangkan noise, dan mengonversi data kategorikal ke dalam format numerik. Selanjutnya dataset dipisahkan menjadi data training dan data testing dengan perbandingan 80:20. Hal ini disebabkan karena algoritma yang digunakan adalah *supervised learning*, yang berarti data terlebih dahulu diajarkan untuk mengenali pola dalam data input sehingga model pengklasifikasi yang dapat digunakan untuk prediksi di masa depan dapat dikembangkan. Terdapat data training sebanyak 12912 data, yang digunakan untuk melatih model serta data testing sebanyak 3228 data, yang digunakan untuk menguji performa model yang telah dilatih. Jumlah data dengan status "Aman" sebanyak 5078, sementara jumlah data dengan status "Waspada" sebanyak 5112 dan jumlah data dengan status "Bahaya" adalah 6082. Data yang telah dibagi, selanjutnya dilatih dengan menggunakan 4 model kernel yang berbeda, yaitu kernel linear, kernel rbf, kernel polynomial dan kernel sigmoid. Dari 4 model kernel SVM tersebut, akan dipilih kernel yang memiliki kinerja terbaik. Setelah memperoleh model dengan tingkat kinerja terbaik, langkah selanjutnya adalah menyimpan model tersebut dan mengintegrasikannya ke dalam alat deteksi banjir yang telah dirancang.

Alat yang digunakan untuk mendeteksi banjir adalah Arduino Uno dan NodeMCU ESP8266. Arduino Uno bertindak sebagai pusat kendali yang mengatur dan mengolah data dari berbagai sensor fisik seperti:

- 1) Sensor ultrasonik (sensor JSN-SR04T), digunakan untuk mengukur jarak antara alat dan objek di depannya yaitu air.
- 2) Sensor DHT22, berfungsi untuk mengukur suhu dan kelembaban udara di sekitar alat deteksi banjir.
- 3) Sensor debit air, untuk mengukur volume dan laju aliran air.
- 4) Sensor curah hujan, yang digunakan untuk mengukur jumlah hujan yang jatuh dalam suatu periode waktu tertentu. Sensor curah hujan bekerja dengan cara mengukur volume air yang terkumpul pada permukaan sensor.

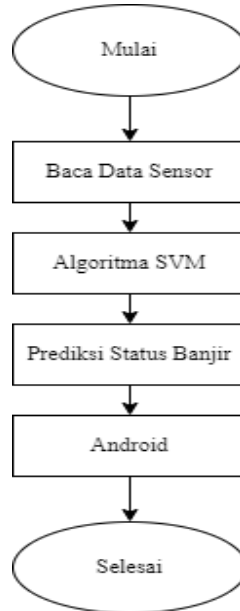
Kemudian, modul Wi-Fi ESP8266, seperti NodeMCU, dapat digunakan untuk menghubungkan sistem ke internet [20]. Pengujian alat dilakukan dengan simulasi buatan. Skema alat yang digunakan adalah sebagai berikut.



Gambar 3. Skema Alat

Adapun tahapan penelitian seperti pada gambar 4. Tahapan penelitian dimulai dengan pembacaan data sensor yaitu sensor ultrasonik untuk mengukur jarak air, sensor *waterflow* untuk mengukur debit air, sensor *rain gauge* untuk mengukur curah hujan dan sensor DHT22

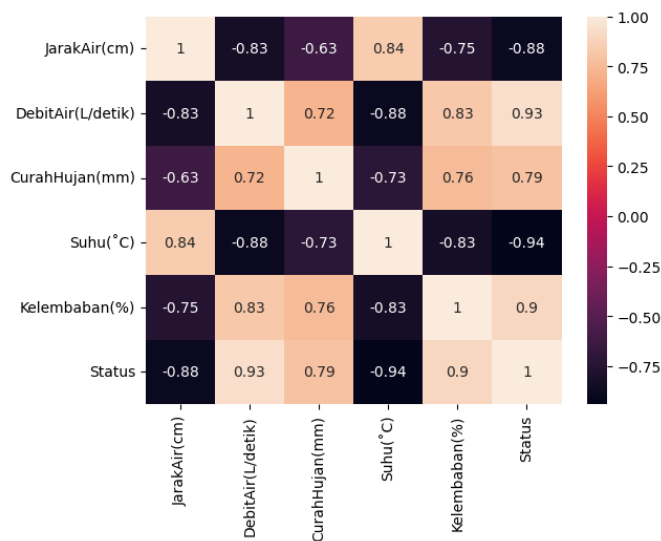
untuk mengukur suhu dan kelembaban. Lalu data hasil pembacaan keempat sensor tersebut akan diolah menggunakan algoritma *Support Vector Machine* (SVM) dan memprediksi status banjir yang didapat. Hasil data sensor dan prediksi status banjir akan dikirim dari database ke aplikasi Android yang telah dirancang untuk ditampilkan.



Gambar 4. Tahapan Penelitian

III. HASIL DAN PEMBAHASAN

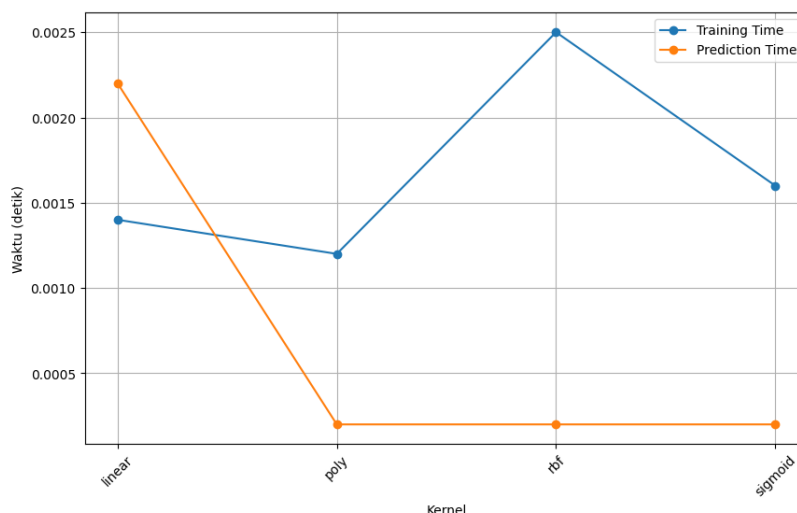
Model kernel *Support Vector Machine* (SVM) yang dapat mengevaluasi model dengan baik, dapat dilihat dari hasil akurasi, presisi, recall, dan F1-score. Pada gambar 5. dilakukan visualisasi antara variabel numerik menggunakan heatmap untuk melihat seberapa eratny hubungan antar variabel. Perhitungan korelasi dan visualisasi heatmap ini dilakukan dengan menggunakan pustaka Pandas dan Seaborn.



Gambar 5. Korelasi Heatmap

Nilai korelasi berada dalam rentang -1 hingga 1, dengan nilai -1 menunjukkan korelasi negatif sempurna, nilai 1 menunjukkan korelasi positif sempurna, dan nilai 0 menunjukkan tidak adanya korelasi. Hasil heatmap pada gambar 5. menunjukkan korelasi negatif moderat

yang berarti bahwa saat salah satu fitur naik, maka fitur lainnya cenderung turun dengan kekuatan yang cukup. Seperti korelasi antara jarak air dan debit air adalah -0.83 , jika jarak air semakin jauh, maka debit airnya cenderung lebih rendah. Begitu pula sebaliknya, jika jarak air semakin tinggi, maka debit airnya cenderung lebih tinggi. Korelasi negatif moderat lainnya ditunjukkan oleh korelasi antara jarak air dan kelembaban dengan nilai korelasi -0.75 serta korelasi antara curah hujan dan suhu dengan nilai korelasi -0.73 . Untuk korelasi negatif sedang, ditunjukkan pada korelasi antara jarak air dan curah hujan dengan nilai korelasinya yaitu -0.63 . Yang berarti bahwa ketika satu fitur naik, fitur lainnya cenderung turun dalam tingkat yang cukup signifikan. Lalu korelasi positif tinggi diantaranya yaitu korelasi antara debit air dan kelembaban dengan nilai korelasi 0.83 dan korelasi antara jarak air dan suhu dengan nilai korelasi 0.84 . Dikatakan korelasi positif tinggi karena saat salah satu fitur naik, fitur lainnya cenderung naik juga dalam tingkat yang signifikan. Korelasi positif moderat ditunjukkan oleh korelasi debit air dan curah hujan dengan nilai korelasi 0.72 dan korelasi antara curah hujan dan kelembaban dengan nilai korelasi 0.76 . Saat dianalisis secara keseluruhan, tidak terdapat pola korelasi yang jelas antara semua fitur yang dipilih. Artinya, hubungan antara fitur-fitur ini tidak mengikuti tren korelasi positif atau negatif yang konsisten.



Gambar 6. Grafik Waktu Pelatihan dan Prediksi SVM

Dari grafik tersebut diperoleh waktu pelatihan (*training time*) dan waktu prediksi (*prediction time*) dari model SVM dengan berbagai jenis kernel yang berbeda. Kernel linear membutuhkan waktu training sekitar 0.0014 detik dan waktu prediction sekitar 0.0022 detik. Untuk kernel polynomial waktu yang diperlukan untuk training sekitar 0.0012 detik dan waktu untuk prediction sekitar 0.0002 detik. Kernel RBF memerlukan waktu training sekitar 0.0025 detik dan waktu prediction sekitar 0.0002 detik. Dan kernel sigmoid memerlukan waktu training sekitar 0.0016 detik dan waktu prediction sekitar 0.0002 detik. Berdasarkan hasil pengukuran waktu tersebut, kernel Polynomial memiliki waktu training dan prediction paling singkat, diikuti oleh kernel Linear, dan kemudian oleh kernel Sigmoid dan RBF.

TABEL IV
METRIK EVALUASI TRAINING

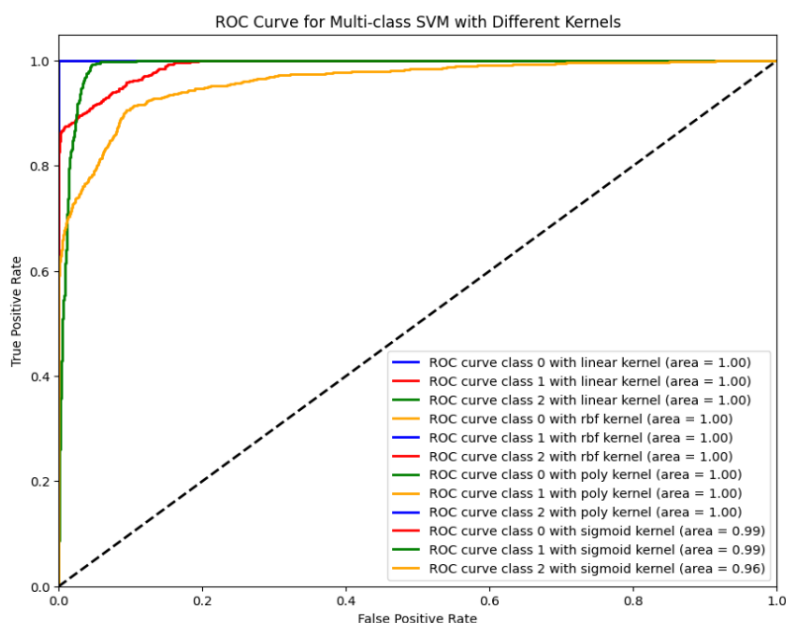
Kernel	Akurasi	Presisi	Recall	F1-Score
Linear	1	1	1	1
Polynomial	1	1	1	1
RBF	1	1	1	1
Sigmoid	0.35	1	0.90	0.95

Pada tabel 3, terlihat SVM dengan kernel linear, kernel RBF, dan kernel polynomial sangat baik dalam melakukan klasifikasi pada dataset ini dengan akurasi dan nilai metrik evaluasi yang tinggi. Namun, SVM dengan kernel sigmoid memiliki performa yang jauh lebih buruk, dengan banyak prediksi yang tidak tepat.

TABEL V
METRIK EVALUASI TESTING

Kernel	Akurasi	Presisi	Recall	F1-Score
Linear	1	1	1	1
Polynomial	1	1	1	1
RBF	1	1	1	1
Sigmoid	0.34	1	0.88	0.94

Pada tabel diatas dapat dilihat bahwa SVM dengan kernel linear, RBF, dan polynomial mencapai akurasi 100% pada dataset yang digunakan. Artinya, ketiga kernel tersebut mampu memisahkan data dengan sempurna untuk tugas klasifikasi tiga kelas yang ada dalam dataset. Metrik lainnya seperti precision, recall, dan f1-score juga mencapai 1.00 untuk setiap kelas, menunjukkan kualitas yang sangat baik. Namun, SVM dengan kernel sigmoid menunjukkan hasil yang jauh lebih rendah dalam hal akurasi (34%) dan metrik lainnya. Hal ini menunjukkan bahwa SVM dengan kernel sigmoid tidak cocok untuk masalah klasifikasi ini dan memberikan prediksi yang kurang akurat. Dalam hal ini, SVM dengan kernel sigmoid mungkin tidak dapat memisahkan data dengan baik dan memprediksi dengan tepat tiga kelas target.



Gambar 7. ROC Curve dengan berbagai kernel

Pada gambar hasil kurva ROC dan nilai AUC di atas, terdapat empat jenis kernel yang diukur yang diindikasikan dengan warna yang berbeda setiap kernelnya. Pada kurva berwarna biru, merah, hijau, dan kuning, ketiganya saling menumpuk dikarenakan memiliki karakteristik kurva ROC yang dan nilai AUC yang sama. Kurva ROC berupa garis lurus yang naik secara vertikal dari (0,0) ke (1,1) tanpa ada lengkungan atau kemiringan. Nilai AUC pada kedua kurva pun memiliki nilai yang sama yaitu 1.00. Ini menunjukkan bahwa model linear, polinomial dan RBF memiliki kemampuan sempurna dalam memprediksi kelas positif dan kelas negatif dengan benar, tanpa kesalahan. Nilai AUC pada model sigmoid sebesar 0.99 menunjukkan bahwa

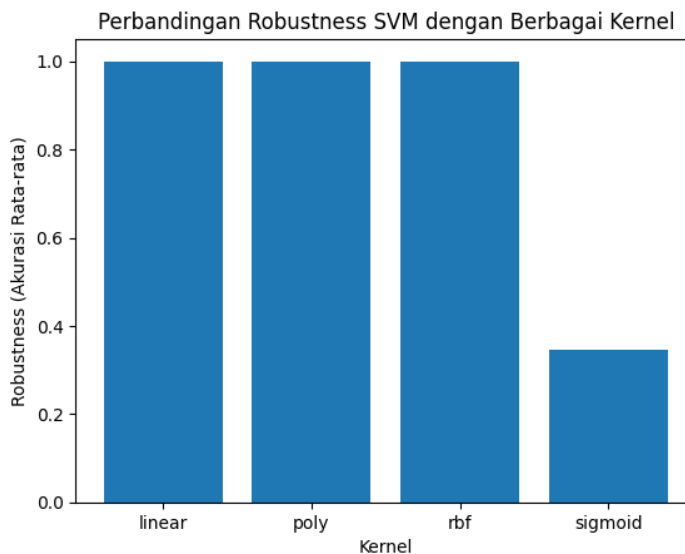
model SVM dengan kernel sigmoid memiliki performa yang baik dalam membedakan kelas 0 dan kelas 1 dari kelas lainnya, tetapi sedikit lebih rendah dibandingkan dengan kernel lainnya untuk kelas pertama (class 0). dan untuk nilai AUC pada model sigmoid kelas 2 sebesar 0.96 yang berarti bahwa memiliki performa yang baik juga tetapi sedikit rendah dibanding dengan kernel lainnya.

Pada tabel VI. merupakan nilai kinerja menggunakan cross-validation. Skor validasi silang merupakan nilai rata-rata akurasi untuk setiap kernel menggunakan 5 lipatan (5-fold-cross-validation).

TABEL VI
NILAI KINERJA MENGGUNAKAN CROSS VALIDATION

Kernel Linear	Kernel Polynomial	Kernel RBF	Kernel Sigmoid
1.0	1.0	1.0	0.35

Tabel diatas menunjukkan bahwa kernel linear, kernel polynomial, dan kernel rbf memiliki nilai kinerja yang sangat tinggi. Hal ini disebabkan oleh penggunaan transformasi polinomial yang memungkinkan model untuk menangkap hubungan non-linear antar fitur. Sehingga kernel ini memiliki stanilitas yang tinggi dan ketahanan yang baik terhadap perubahan dan variasi. Sedangkan, kernel sigmoid jauh lebih rendah yaitu hanya sekitar 35%. Hal ini mengindikasikan bahwa kernel sigmoid mungkin tidak cocok dengan dataset atau jenis masalah seperti ini dan mungkin menghadapi kesulitan dalam membedakan kelas-kelas dalam dataset ini.



Gambar 8. Perbandingan Robustness SVM dengan Berbagai Kernel

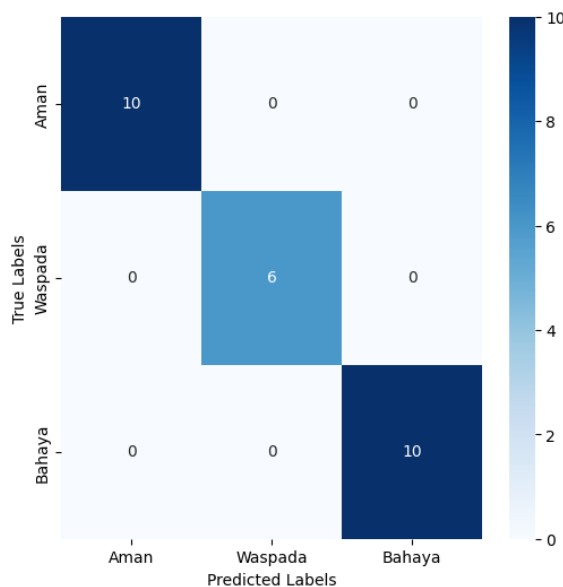
Pada gambar 8. dipaparkan perbandingan kinerja SVM dengan berbagai jenis kernel dalam situasi keseluruhan. Hasil analisis tersebut mengindikasikan bahwa kernel linear, RBF, dan polynomial menunjukkan hasil yang lebih unggul dalam menangani pola-pola yang kompleks, serta mampu menghasilkan tingkat ketahanan (robustness) yang lebih tinggi. Namun, kernel sigmoid menunjukkan performa yang rendah dalam dataset yang diberikan. Setelah memeriksa berbagai kernel, kernel polynomial menunjukkan performa terbaik dalam hal waktu pelatihan dan prediksi, akurasi, presisi, recall, kurva ROC dan F1-score. Berdasarkan hasil ini, kernel polynomial dipilih sebagai model SVM untuk memprediksi status banjir.

Setelah berhasil memperoleh model SVM dengan kinerja perfoma tertinggi, model tersebut akan diimplementasikan pada alat deteksi banjir. Dilakukan pengambilan data dari sensor dengan total 131 data uji. Hasil dari data yang diperoleh kemudian disajikan dalam Tabel VII.

TABEL VII
HASIL TESTING MENGGUNAKAN ALAT DETEKSI BANJIR

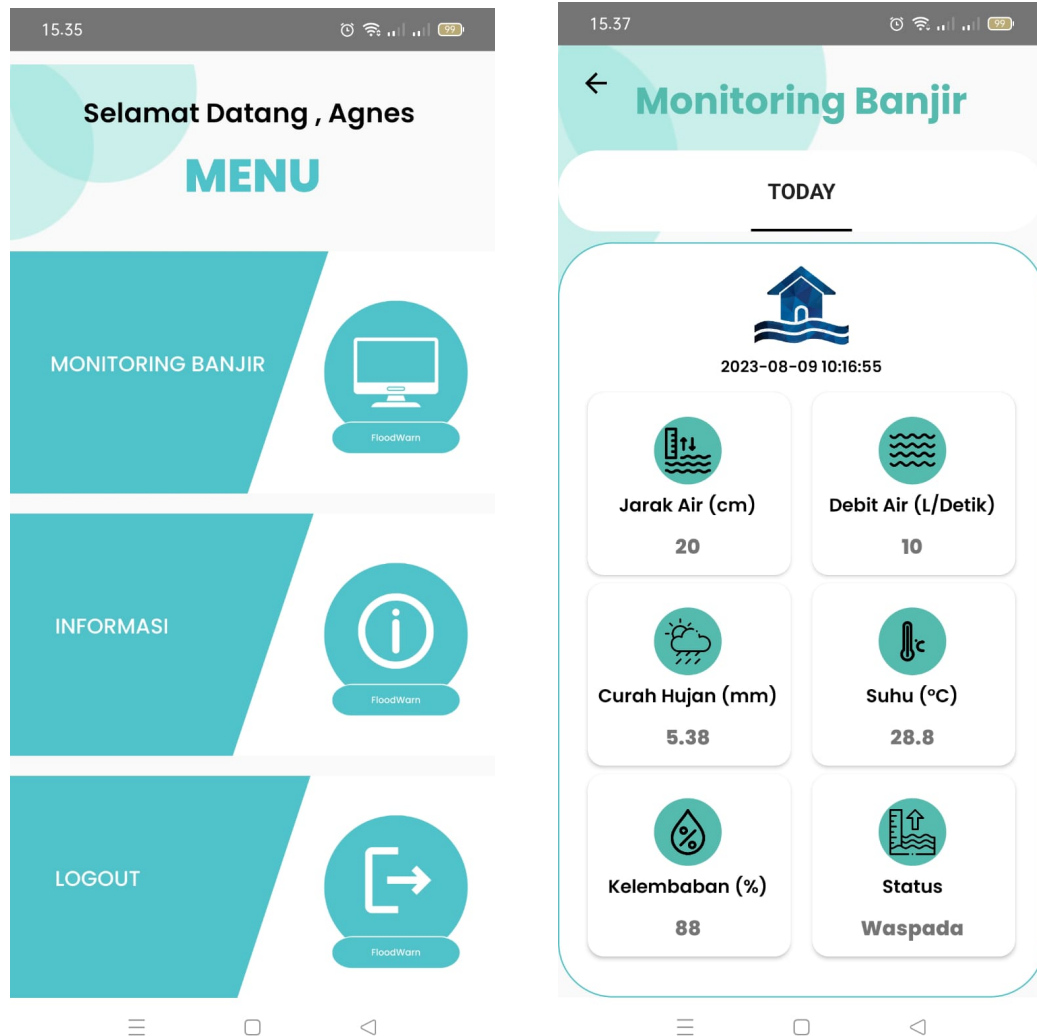
No.	Jarak Air (cm)	Debit Air (L/detik)	Curah Hujan (mm)	Suhu (°C)	Kelembaban (%)	Hasil Prediksi Status
1	28	15	2.69	30.60	67	Aman
2	29	22	1.35	30.30	68.70	Aman
3	34	15	3.37	30.70	69.20	Aman
4	26	10	0.27	31	67	Aman
5	18	10	1.2	30.04	65	Aman
6	27	36	2.69	30.80	67	Waspada
7	30	44	2.69	31.10	67.70	Waspada
8	37	6	37.69	31.20	68.80	Waspada
9	25	31	2.69	30.70	67.00	Waspada
10	20	13	1.35	30	86.5	Waspada
11	34	158	5.38	31.40	67.30	Bahaya
12	32	11	232.86	31	68.30	Bahaya
13	25	51	37.69	31	67.10	Bahaya
14	183	20	344.58	31	68.20	Bahaya
15	171	224	152.1	30.6	69.3	Bahaya

Berdasarkan hasil uji dari alat deteksi banjir, diperoleh hasil klasifikasi untuk 131 percobaan data yang telah dijalankan. Akurasi untuk setiap label dapat dihitung dengan memanfaatkan *confusion matrix*, sebagaimana terlihat pada gambar 9.



Gambar 9. Hasil Confusion Matrix Testing

Hasil confusion matrix diatas menunjukkan bahwa ada 10 data yang sebenarnya termasuk dalam kelas "Aman" dan juga diprediksi dengan benar oleh model sebagai "Aman". Ada 6 data yang sebenarnya termasuk dalam kelas "Waspada" dan juga diprediksi dengan benar oleh model sebagai "Waspada." Dan da 10 data yang sebenarnya termasuk dalam kelas "Bahaya" dan juga diprediksi dengan benar oleh model sebagai "Bahaya." Secara keseluruhan, model yang menggunakan kernel polinomial telah mengklasifikasikan setiap sampel dengan benar sesuai dengan kelas sebenarnya pada dataset uji.



Gambar 10. (a) Tampilan Menu Utama (b) Tampilan Menu Monitoring Banjir

Adapun bentuk tampilan pada aplikasi android yang menampilkan hasil data sensor dan hasil klasifikasi dari data yang telah diambil. Gambar 10(a) menggambarkan pilihan menu dalam aplikasi pendeteksi banjir. Pilihan menu ini terdiri dari tiga menu utama, yakni

- "Monitoring Banjir" yang memungkinkan pemantauan status banjir.
- "Informasi" yang berisi penjelasan tentang aplikasi.
- "Logout" yang berfungsi untuk keluar dari aplikasi.

Gambar 10(b) merupakan tampilan menu "Monitoring Banjir" yang terdiri dari hasil data sensor alat dan hasil deteksi. Database akan mengirimkan hasil data sensor dan hasil deteksi tersebut ke aplikasi dan ditampilkan pada menu ini. Database yang digunakan adalah MySQL. Menu monitoring banjir, menampilkan data monitoring yang akan diperbarui setiap 1 menit.

IV. KESIMPULAN

Implementasi algoritma Support Vector Machine (SVM) dalam deteksi banjir telah membuktikan kemampuannya dalam memprediksi status banjir dengan performa yang sangat tinggi. Dalam penelitian ini, berbagai jenis kernel seperti linear, polynomial, RBF, dan sigmoid dievaluasi untuk menemukan kernel terbaik yang cocok untuk mendeteksi status banjir. Berdasarkan hasil evaluasi algoritma, kernel polinomial menghasilkan performansi terbaik diantara kernel lainnya. Pada hasil akurasi pelatihan menghasilkan nilai 1.0 dan akurasi pengujian 1.0. Hasil perbandingan waktu dan skabilitas SVM berbagai kernel disimpulkan

kernel polinomial memiliki waktu training dan testing paling singkat dengan waktu 0.0012s dan 0.0002s. Berdasarkan hasil confusion matrix training dan testing, kernel polinomial menghasilkan nilai yang tetap yaitu 1.0. Sensitivitas (recall), presisi, dan F1-Score yang dihasilkan sebesar 1.0. Ini menunjukkan bahwa model secara konsisten dapat membedakan dengan baik antara kelas-kelas yang ada dalam data. Pada hasil parameter nilai AUC dan kurva ROC, model polinomial memiliki kemampuan yang baik dalam memprediksi kelas positif dan kelas negatif dengan benar, tanpa kesalahan. Parameter robustness pada kernel polinomial memiliki nilai yang paling tinggi yaitu 1.0. Hal ini menjadikan kernel Polynomial sebagai pilihan optimal untuk diimplementasikan pada perangkat pendeteksi banjir. Pada pengujian menggunakan alat deteksi banjir, hasil confusion matrix menunjukkan bahwa model telah mengklasifikasikan setiap sampel dengan benar sesuai dengan kelas sebenarnya pada dataset uji.

REFERENSI

- [1] Risdiandi R. Analisis Cara Kerja Sensor Ultrasonik Menggunakan Mikrokontroler Arduino Uno Untuk Merancang Alat Deteksi Banjir Secara Otomatis. OSF Preprints. Jan. 2021.
- [2] Annur CM. Banjir sampai Kekeringan, Ini Bencana Alam di Indonesia pada 2022. Databoks. Jan. 04, 2023. Available: <https://databoks.katadata.co.id/datapublish/2023/01/04/banjir-sampai-kekeringan-ini-bencana-alam-di-indonesia-pada-2022>.
- [3] Akbar H, Aryani D, Ulum MB. Deteksi Banjir Area Perkotaan Berbasis Citra Digital Convolutional Neural Network (VGG19). Jurnal Teknik Mesin, Elektro dan Ilmu Komputer. 2022; 2(3): 82-91.
- [4] Tarigan J, Betan AD. Sistem Perancangan Pendeteksi Banjir Secara Dini Menggunakan Mikrokontroler Arduino Uno. Jurnal Teknik Mesin. 2019; 2(2): 63-67.
- [5] Tenggono A, Wijaya Y, Kusuma E, et al. Sistem Monitoring dan Peringatan Ketinggian Air berbasis Web dan SMS Gateway. Sisfotenika. 2015; 5(2): 119-129.
- [6] Purnamasari D, Jonathan JH, Perdana SY, Ihsani F, Wicaksana SIWW. Get Easy Using Weka. Jakarta Timur: Dapur Buku. 2013.
- [7] Wahyudi Y, Faiqurahman M, Risqiwati D. Sistem Iot Untuk Deteksi Bencana Banjir Menggunakan Algoritma C4.5 Dan Modul Komunikasi ESP 8266. In Prosiding SENTRA (Seminar Teknologi dan Rekayasa). 2021; 6: 295-303.
- [8] Mauliana P. Prediksi Banjir Sungai Citarum dengan Logika Fuzzy Hasil Algoritma Particle Swarm Optimization. Jurnal Informatika. 2016; 3(2).
- [9] Stanley F, Lisangan EA. Sistem Dan Simulasi Deteksi Banjir Untuk Peringatan Dini Diolah Memakai Metode KNN Berbasis Arduino. TEMATIKA: Jurnal Penelitian Teknik Informatika dan Sistem Informasi. 2020; pp. 9-22.
- [10] Cahyo MPD. Kinerja Algoritma Support Vector Machine dalam Menentukan Kebenaran Informasi Banjir di Twitter (Doctoral dissertation, Universitas Negeri Jakarta). 2017.
- [11] Ritonga AS, Purwaningsih ES. Penerapan Metode Support Vector Machine (SVM) Dalam Klasifikasi Kualitas Pengelasan Smaw (Shield Metal Arc Welding). Jurnal Ilmiah Edutic: Pendidikan dan Informatika. 2018; 5(1): 17-25.
- [12] Faiza IM, Andriani W. Tinjauan Pustaka Sistematis: Penerapan Metode Machine Learning untuk Deteksi Bencana Banjir. Jurnal Minfo Polgan. 2022; 11(2): 59-63.
- [13] Ritonga AS, Purwaningsih ES. Penerapan Metode Support Vector Machine (SVM) Dalam Klasifikasi Kualitas Pengelasan Smaw (Shield Metal Arc Welding). Jurnal Ilmiah Edutic: Pendidikan dan Informatika. 2018; 5(1): 17-25.
- [14] Budiharto W, Prabawati TA. Machine Learning dan Computational Intelligence. Yogyakarta: Penerbit Andi; 2016.

- [15] Parapat IM, Furqon MT, Sutrisno S. Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. 2018; 2(10): 3163-3169.
- [16] Praghakusma AZ. Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus: Komisi Pemberantasan Korupsi (KPK)). Doctoral dissertation, UPN Veteran Yogyakarta; 2021.
- [17] Sasongko TB. Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Klasifikasi Jalur Minat SMA). *Jurnal Teknik Informatika dan Sistem Informasi*. 2016; 2(2).
- [18] J. M. Ashfaqe and F. Markowet, "Introduction to Support Vector Machines and Kernel Methods," no. April, pp. 1–9, 2019.
- [19] Praghakusma AZ. Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus: Komisi Pemberantasan Korupsi (KPK)). Doctoral dissertation, UPN Veteran Yogyakarta; 2021.
- [20] Samsugi S. *Internet of Things (IoT): Sistem Kendali Jarak Jauh Berbasis Arduino dan Modul WiFi Esp8266*. ReTII. 2017.