

BAB I

PENDAHULUAN

1.1. Latar Belakang

Autisme merupakan salah satu gangguan perkembangan otak yang menyebabkan kelainan dalam proses tumbuh kembang individu. Penderita autisme biasanya menunjukkan berbagai gejala, seperti gangguan pada kemampuan kognitif, perilaku, bahasa, komunikasi, serta interaksi sosial. Individu penyandang autisme umumnya menunjukkan perilaku yang cenderung memberontak dan sering melakukan aktivitas tertentu secara berulang-ulang, berbeda dengan individu pada umumnya yang menunjukkan pola perilaku yang lebih bervariasi [1]. Tingkat ekspresi dari setiap individu dapat sangat beragam dalam spektrum yang luas. Saat ini, para peneliti seperti Apeksha et al. (2023) dan Patel et al. (2023) tengah berfokus pada upaya mempelajari dan mengkategorikan pola bahasa yang muncul pada individu dengan autisme. Penelitian-penelitian ini berusaha untuk mengidentifikasi ciri-ciri khusus pada individu dengan autisme, serta untuk menggali pemahaman lebih dalam mengenai mekanisme neurologis yang mendasari kondisi tersebut. Dengan memahami pola-pola bahasa ini, diharapkan dapat meningkatkan diagnosis dini dan pemahaman tentang keberagaman gejala yang ditunjukkan oleh individu dengan autism [2][3].

Dalam penelitian yang berjudul *Investigating Multiclass Autism Spectrum Disorder Classification Using Machine Learning Techniques* oleh Bawa et al. (2024), upaya dilakukan untuk membangun sistem klasifikasi guna membedakan anak-anak, remaja, dan dewasa dengan ASD. Deteksi dini menjadi penting untuk mengurangi gejala ASD dan meningkatkan kualitas hidup, namun protokol pemeriksaan medis yang ada masih terbatas. Oleh karena itu, diperlukan teknik otomatis untuk mengklasifikasikan individu dengan ASD. Pada populasi dewasa, gejala ASD sering tumpang tindih dengan kondisi kesehatan mental lainnya, sehingga hanya fokus pada pemindaian otak atau pola bicara mungkin tidak cukup untuk mengevaluasi kondisi secara menyeluruh. Tes skrining perkembangan dapat memberikan evaluasi yang lebih komprehensif untuk deteksi dini ASD pada ketiga kelompok usia. Penelitian ini juga mengeksplorasi berbagai teknik klasifikasi seperti Naive Bayes, Decision Tree, Random Forest, K-Nearest Neighbor, Support Vector Machine, dan Logistic Regression, serta menggunakan

klasifikasi multi-kelas untuk diagnosis ASD yang lebih akurat. Hasil penelitian menunjukkan bahwa penerapan teknik pembelajaran mesin, khususnya dengan hyperparameterization menggunakan Grid Search, meningkatkan akurasi model dalam klasifikasi ASD, dengan SVM mencapai telah mencapai akurasi yang cukup tinggi hingga 99,55%. Penyesuaian parameter yang tepat untuk dataset yang tidak seimbang membantu mencegah overfitting dan meningkatkan generalisasi model. Meskipun model KNN dan Naïve Bayes tidak menunjukkan perubahan signifikan, temuan ini menekankan pentingnya pembelajaran mesin dalam deteksi dini ASD, meski distribusi label yang tidak merata dapat menyebabkan bias pada model [4]

Berdasarkan hasil penelitian diatas, belum terdapat penelitian lanjutan terkait dengan optimasi dari model *machine learning* yang terbaik dari hasil penelitian untuk mendapatkan akurasi yang lebih tinggi untuk data terkait. SVM merupakan model yang memiliki akurasi paling tinggi diantara model lainnya pada hasil yang didapatkan pada penelitian. Namun demikian model SVM membutuhkan optimasi lebih lanjut agar didapatkan peningkatan akurasi, salah satu teknik peningkatannya adalah menggunakan *Principal Component Analysis* (PCA) untuk mereduksi fitur dari data yang akan digunakan dalam klasifikasi nantinya pada SVM. Selain itu penentuan dalam parameter yang digunakan untuk proses SVM juga cukup penting untuk peningkatan dalam akurasi. Penelitian sebelumnya peneliti telah menggunakan Grid Search sebagai hyperparameterization untuk peningkatan SVM. Namun demikian terdapat teknik optimasi parameter lain yang yakni *Particle Swarm Optimization* (PSO) yang memiliki pendekatan yang berbeda dengan menggerakkan partikel-partikel dalam ruang parameter untuk mencari solusi optimal. PSO yang kemungkinan besar lebih efisien dalam mengeksplorasi ruang parameter besar, menjadikannya alternatif yang lebih cepat dan efektif, terutama pada ruang parameter yang kompleks.

Dalam penelitian yang dilakukan oleh Xiu Chen et al. (2021) berjudul "*Identification of architectural elements based on SVM with PCA: A case study of sandy braided river reservoir in the Lamadian Oilfield, Songliao Basin, NE China*", penerapan PCA untuk mereduksi dimensi data terbukti meningkatkan akurasi identifikasi pada algoritma SVM. Hasil penelitian tersebut menunjukkan bahwa akurasi identifikasi parameter fitur logging setelah reduksi dimensi berbasis PCA mencapai 92,08%, yang lebih tinggi 6,93 persen dibandingkan dengan akurasi model SVM yang menggunakan parameter fitur tanpa reduksi dimensi [5]. Temuan ini sejalan dengan hasil penelitian ini

yang juga menunjukkan peningkatan kinerja model SVM melalui penerapan teknik reduksi dimensi seperti PCA.

Dalam penelitian yang dilakukan oleh Samantaraya et al. (2023) berjudul *"Prediction of Flood Discharge Using Hybrid PSO-SVM Algorithm in Barak River Basin"*, ditemukan bahwa penggunaan kombinasi PSO dengan SVM dapat menghasilkan model yang lebih efektif dalam memprediksi curah hujan dan aliran banjir bulanan. Model PSO-SVM menunjukkan kinerja yang lebih baik dibandingkan dengan model konvensional lainnya, dengan nilai RMSE sebesar 0.04962, R^2 0.98918, dan NSE 0.99334 untuk dua stasiun pengukuran[6]. Penelitian ini menunjukkan bahwa teknik hybrid PSO-SVM dapat menjadi alternatif yang lebih unggul dalam pemodelan dan dapat diterapkan pada berbagai skenario pemodelan lainnya.

Berdasarkan penelitian-penelitian sebelumnya, penulis ingin melakukan perbandingan performa model Support Vector Machine (SVM) dengan menggunakan beberapa metode optimasi, yaitu Principal Component Analysis (PCA), Particle Swarm Optimization (PSO), dan GridSearch, untuk mengklasifikasi Autism Spectrum Disorder (ASD). Data yang digunakan adalah data skrining autisme untuk anak-anak, remaja, dan dewasa. Perbandingan yang dilakukan akan mempertimbangkan faktor-faktor penting seperti accuracy, precision, recall, F1-Score, dan AUC. Penelitian ini bertujuan untuk mengevaluasi teknik mana yang paling efektif dalam meningkatkan kinerja SVM untuk klasifikasi ASD dan memberikan wawasan tentang penggunaan kombinasi metode optimasi dalam konteks tersebut.

1.2. Rumusan Masalah

Dari uraian latar belakang yang telah disajikan, dapat diidentifikasi rumusan masalah yang menjadi fokus pembahasan sebagai berikut:

1. Bagaimana pengaruh PCA dalam mereduksi dimensi data terhadap performa klasifikasi Autism Spectrum Disorder menggunakan SVM?
2. Bagaimana perbandingan efektivitas PSO dan Grid Search sebagai metode optimasi parameter dalam meningkatkan kinerja model SVM untuk klasifikasi Autism Spectrum Disorder?
3. Kombinasi metode PCA, PSO, dan Grid Search mana yang memberikan hasil terbaik dalam klasifikasi Autism Spectrum Disorder?

4. Bagaimana dampak optimasi parameter menggunakan PSO dan Grid Search terhadap efisiensi waktu pemrosesan model SVM untuk klasifikasi Autism Spectrum Disorder?

1.3. Tujuan Penelitian

Merujuk pada rumusan masalah diatas, adapun tujuan utama yang ingin dicapai penulis dari penelitian ini dibedakan menjadi 2 tujuan yaitu, tujuan umum dan tujuan khusus:

1.3.1. Tujuan Umum

Penelitian ini secara umum bertujuan untuk menganalisis pengaruh reduksi dimensi data menggunakan PCA terhadap kinerja model SVM serta mengevaluasi dan membandingkan efektivitas metode optimasi parameter PSO dan Grid Search dalam meningkatkan akurasi klasifikasi Autism Spectrum Disorder (ASD).

1.3.2. Tujuan Khusus

Tujuan khusus dalam penelitian ini adalah sebagai berikut:

- a. Menganalisis pengaruh penerapan PCA dalam mereduksi dimensi data terhadap performa klasifikasi Autism Spectrum Disorder menggunakan SVM.
- b. Membandingkan efektivitas PSO dan Grid Search sebagai metode optimasi parameter dalam meningkatkan kinerja model SVM untuk klasifikasi Autism Spectrum Disorder.
- c. Menentukan kombinasi metode PCA, PSO, dan Grid Search yang memberikan hasil terbaik dalam klasifikasi Autism Spectrum Disorder.
- d. Mengkaji dampak optimasi parameter menggunakan PSO dan Grid Search terhadap efisiensi waktu pemrosesan model SVM untuk klasifikasi Autism Spectrum Disorder.

1.4. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Bagi Akademisi

Penelitian ini diharapkan dapat memberikan kontribusi penting dalam literatur terkait penerapan PCA untuk reduksi dimensi data dan perbandingan antara PSO dan Grid Search sebagai metode optimasi parameter pada model SVM. Hasil

penelitian ini dapat menjadi referensi bagi pengembangan metode machine learning di masa depan, khususnya dalam konteks klasifikasi Autism Spectrum Disorder. Penelitian ini juga dapat membuka peluang untuk penelitian lebih lanjut, terutama dalam eksplorasi kombinasi metode pembelajaran mesin yang lebih inovatif dan efisien, serta aplikasi metode ini pada kasus klasifikasi lainnya.

2. Bagi Praktisi

Bagi praktisi, penelitian ini diharapkan dapat memberikan wawasan dan pemahaman tentang penerapan teknik-teknik pembelajaran mesin, seperti PCA, PSO, dan Grid Search, untuk meningkatkan kinerja model SVM dalam klasifikasi Autism Spectrum Disorder. Hasil penelitian ini dapat digunakan untuk mengoptimalkan proses klasifikasi data dengan teknik-teknik tersebut, memberikan solusi yang lebih efisien dan akurat dalam pengolahan data berdimensi tinggi. Selain itu, temuan ini dapat menjadi acuan bagi praktisi dalam mengembangkan dan mengimplementasikan metode pembelajaran mesin untuk aplikasi serupa di bidang lain.

1.5. Batasan Masalah

Bagian dari uraian latar belakang yang telah disajikan, dapat diidentifikasi rumusan masalah yang menjadi fokus pembahasan sebagai berikut:

- a. Data yang digunakan dalam penelitian merupakan data sekunder (data yang diambil secara tidak langsung oleh penulis) yang diakses melalui *website UC Irvine Machine Learning Repository*
- b. Data yang digunakan adalah data skrining untuk 3 kelas umur, anak-anak, remaja dan dewasa.
- c. Peneliti menggunakan data yang memiliki 20 fitur didalamnya, yang terdiri dari antara lain seperti usia, jenis kelamin, etnisitas, kondisi kelahiran, dan hasil jawaban pertanyaan terkait autisme.
- d. Luaran yang dihasilkan adalah *accuracy, precision, recall, F1-Score, area under the curve* dan waktu komputasi dari masing – masing skenario pengujian yang ditentukan.

