

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gangguan spektrum autisme atau *autism spectrum disorder* (ASD) adalah gangguan perkembangan pada manusia terutama dalam berperilaku, yang secara umum disebabkan oleh kelainan struktur otak atau fungsi otak (Milyawati dan Hastuti, 2009). ASD perlu diprediksi dengan akurat karena waktu yang dibutuhkan untuk diagnosis (*screening*) sangat panjang dan prosedur yang dilakukan membutuhkan banyak biaya (Thabtah, 2017). ASD tidak hanya terdeteksi pada usia anak-anak, tetapi dapat terjadi ketika remaja dan dewasa atau seiring bertambahnya usia. Berbagai penyebab hal tersebut antara lain: pola perilaku, emosi yang naik-turun, kesulitan beradaptasi, hingga depresi. Faktor penyebab ini berkaitan erat dengan sejauh mana tingkat kecerdasan dan keterampilan seseorang untuk berkomunikasi dengan lingkungan sekitarnya (Fajar dan Firdaus, 2017).

Panjangnya prosedur dan mahalnya biaya diagnosis ASD dapat diatasi dengan mengurangi atribut yang terlibat dalam kuesioner diagnosis ASD seperti yang telah dilakukan pada (Thabtah, 2017). Penelitian tersebut secara khusus menekankan parameter yang terdapat pada metode *Diagnostic and Statistical Manual* (DSM) sebagai urgensi dalam memperbarui alat skrining klinis untuk mencerminkan perubahan yang diusulkan dalam manual DSM-5 dan berhasil menemukan parameter penentu (kriteria baru) sebagai gejala gangguan spektrum autisme dalam metode DSM-5.

Penelitian lain mengenai prediksi ASD (Thabtah, dkk., 2018) mengungkapkan bahwa perlu dilakukan prediksi dini mendahului terdeteksinya gejala ASD melalui pemilihan parameter menggunakan metode *variable analysis* (VA). VA terbukti dapat mengurangi gejala gangguan spektrum autisme pada remaja berdasarkan DSM-5 yang awalnya berjumlah dua puluh menjadi delapan, namun tetap mempertahankan akurasi prediksi kompetitif, sensitivitas, dan tingkat spesifisitas.

Berbagai teknik *machine learning* untuk menghitung akurasi prediksi telah diterapkan oleh peneliti lain, diantaranya *neural network* (NN) (Mohammad, dkk., 2014) dan *Support Vector Machine* (SVM) (Platt, 1998). NN memiliki kelebihan pada prediksi nonlinear, kuat di parallel processing dan kemampuan mentoleransi kesalahan tinggi, tapi memiliki kelemahan pada perlunya data latih yang besar, *over-fitting*, lambatnya konvergensi, dan sifatnya yang *local optimum* (Capparuccia, dkk., 2007). SVM dapat memecahkan masalah NN, yaitu *over-fitting*, lambatnya konvergensi, dan sedikitnya data latih (Vapnik, 1995). Pernyataan tersebut sesuai dengan karakteristik data ASD pada remaja dalam penelitian ini yang datanya sedikit. Tetapi SVM memiliki kelemahan pada sulitnya pemilihan parameter SVM yang optimal (Coussement dan Van den Poel, 2008). Beberapa algoritma direkomendasikan oleh peneliti dunia untuk mengoptimasi parameter pada SVM, seperti: *simulated annealing* (SA) (Jia, dkk., 2011), *particle swarm optimization* (PSO) (Wang, 2014), *grid search* (Naufal, dkk., 2015), dan *genetic algorithm* (GA) (Guo, dkk., 2012).

Simulated annealing efektif pada pemuatan masalah optimasi pola namun memiliki kecenderungan untuk terjebak dalam minimum lokal (Zameer, dkk., 2014). Algoritma PSO memiliki kesulitan menentukan nilai yang optimum dalam mengoptimasi lebih dari sepuluh parameter, begitu juga dengan *grid search* (Bergstra dan Bengio, 2012). GA dapat mengatasi masalah yang nonlinier dengan diskontinuitas dan minimal lokal secara efisien, serta lebih efisien dalam mengoptimasi lebih dari sepuluh parameter (Machairas, dkk., 2014). Pada penelitian ini algoritma genetika akan diterapkan untuk pemilihan parameter SVM yang sesuai dan optimal, sehingga hasil prediksi gangguan spektrum autisme pada remaja lebih akurat.

1.2 Identifikasi Masalah

Support Vector Machine adalah teknik klasifikasi yang memiliki performa sangat baik untuk prediksi rentet waktu karena dapat memecahkan masalah *overfitting*, lambatnya konvergensi, dan sedikitnya data latih. Di balik performanya yang sangat baik; SVM memiliki kelemahan terkait sulitnya melakukan pemilihan parameter yang optimal karena dilakukan secara *trial and error*, sehingga menyebabkan tingkat akurasi prediksi menjadi rendah. GA dapat diimplementasikan untuk mengatasi kelemahan dalam pemilihan parameter yang optimal pada SVM untuk meningkatkan akurasi prediksi.

1.3 Pembatasan Masalah

Batasan masalah pada penelitian tesis ini dapat adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah *public dataset* “*Autistic Spectrum Disorder Screening Data for Adolescent*” oleh Fadi Thabtah, diunduh dari *UCI repository* (<http://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++>).
2. Data set hanya dapat digunakan untuk memprediksi usia remaja yang berumur dua belas sampai dengan enam belas tahun.
3. Prediksi menghasilkan dua kelas, yaitu: ya (positif mengalami gangguan spektrum autisme) dan tidak (negatif mengalami gangguan spektrum autisme).

1.4 Rumusan Masalah

Berdasarkan uraian dari latar belakang permasalahan di atas didapatkan rumusan masalah, yaitu bagaimana implementasi dan akurasi metode SVM yang dioptimasi melalui penerapan GA dalam prediksi ASD pada remaja?

1.5 Tujuan Penelitian

Tujuan dari penelitian tesis ini, yaitu melakukan klasifikasi data ASD pada remaja menggunakan metode SVM dan membandingkan akurasinya dengan SVM yang parameternya dioptimasi menggunakan GA. SVM yang dioptimasi melalui penerapan GA bertujuan untuk meningkatkan akurasi prediksi ASD pada remaja

sehingga dapat berkontribusi ke bidang ilmu pengetahuan mengenai signifikansi parameter SVM yang perlu dioptimalkan salah satunya dengan menggunakan GA.

1.6 Manfaat Penelitian

Manfaat yang diharapkan dari seluruh rangkaian kegiatan penelitian serta hasil penelitian tesis ini adalah sebagai berikut:

1. Manfaat praktis yang diharapkan adalah bahwa seluruh tahapan penelitian serta hasil penelitian yang diperoleh dapat memperluas wawasan dan sekaligus memperoleh pengetahuan empirik mengenai penerapan GA untuk optimasi parameter SVM dalam prediksi ASD pada remaja. Bagi pihak-pihak yang berkepentingan dengan hasil penelitian, penulis berharap manfaat hasil penelitian dapat diterima sebagai kontribusi untuk membantu diagnosis pada remaja dan untuk mengurangi mahalanya biaya diagnosis ASD, dimana hal tersebut memerlukan waktu yang lama.
2. Manfaat akademis yang diharapkan adalah bahwa hasil penelitian dapat dijadikan rujukan bagi upaya pengembangan ilmu komputer dan berguna sebagai referensi bagi mahasiswa yang melakukan kajian terhadap metode *data mining* khususnya klasifikasi menggunakan SVM dengan parameter yang dioptimasi.