

# BAB I

## PENDAHULUAN

### A. Latar Belakang

*Stroke* merupakan kondisi yang terjadi ketika asupan darah ke otak terganggu atau berkurang akibat penyumbatan (*stroke* iskemik) atau pecahnya pembuluh darah (*stroke* hemoragik), sehingga jaringan otak tidak mendapatkan oksigen dan nutrisi (Setyopranoto, 2003). *Stroke* adalah keadaan darurat medis karena sel-sel otak mati dalam beberapa menit. *Stroke* merupakan penyakit serius dan membahayakan nyawa, oleh sebab itu harus ditangani secepatnya. Tindakan penanganan yang cepat dan tepat dapat meminimalkan tingkat kerusakan otak dan mencegah kemungkinan munculnya komplikasi lainnya (Kurniawan Nugroho *et al.*, 2022).

*Stroke* merupakan penyebab kematian nomor tiga di dunia. Di rumah sakit, *stroke* merupakan penyebab kematian kedua di dunia setelah penyakit jantung coroner (Saraswati, D, 2021). Hasil Riset Kesehatan Dasar (Rikesdas) 2013 oleh Kementerian Kesehatan (Kemenkes) RI menunjukkan bahwa prevalensi *stroke* di Indonesia adalah 12,1/1000 orang. *Stroke* merupakan penyebab kematian utama di Indonesia. Lebih banyak pria (7,1%) mengalami *stroke* dibandingkan wanita (6,8%) (Rikesdas, 2013).

*Stroke* sering terjadi pada usia lanjut, sehingga usia merupakan faktor terpenting penyebab terjadinya *stroke* (Setyopranoto, 2003). *Stroke* dapat menyebabkan kelumpuhan secara tiba-tiba di salah satu sisi wajah, lengan atau tungkai. *Stroke* juga dapat menyebabkan gangguan penglihatan, kesulitan dalam berbicara dan memahami bahasa (Afasia), kesadaran yang menurun serta gangguan fungsi intelektual (demensia).

Berdasarkan hasil Riset Kesehatan Dasar (Rikesdas) maka penelitian ini akan melakukan klasifikasi terhadap penyakit *stroke*. Dilakukan klasifikasi yaitu untuk mempermudah pengolahan data sehingga analisis juga akan

cepat. Pentingnya klasifikasi *stroke* pada penelitian ini untuk meningkatkan akurasi diagnosa menggunakan fitur-fitur pada dataset yang digunakan.

Penelitian klasifikasi penyakit *stroke* ini, penulis menggunakan *machine learning*. Sebagaimana *machine learning* dilatih untuk membuat klasifikasi atau prediksi dalam pengembangan data. Menurut (Hania, 2017) *machine learning* sering digunakan untuk menggantikan perilaku manusia dalam menyelesaikan masalah. Salah satu metode *machine learning* untuk klasifikasi yang terkenal adalah metode *Support Vector Machine* (SVM) (Ritonga & Purwaningsih, 2018).

*Support Vector Machine* (SVM) adalah algoritma yang sering digunakan untuk suatu klasifikasi yang menggunakan data latih untuk memprediksi keanggotaan data pengamatan dan menggunakan data uji untuk validasi model (Mukti, 2020). SVM bisa menghasilkan *accuracy* tinggi dan memberi kesalahan yang relatif kecil (Kusumaningrum, 2017). Metode SVM sudah banyak digunakan dengan berbagai topik, seperti pada penelitian (Kusumaningrum, 2017) yang berjudul “Optimasi Parameter *Support Vector Machine* Menggunakan *Genetic Algorithm* untuk Klasifikasi Microarray Data”, pada penelitiannya menghasilkan *accuracy* 75%.

Proses klasifikasi pada model SVM menggunakan kernel *Radial Basic Function* (RBF), karena kernel RBF dapat mengatasi permasalahan nonlinearitas pada data serta menghasilkan performa yang baik (Kusumaningrum, 2017). Parameter-parameter terbaik pada SVM seperti C dan  $\gamma$  harus dilakukan optimasi parameter pada model SVM. Optimasi parameter tersebut dilakukan untuk memperoleh *performance* yang optimal. Metode optimasi parameter yang dapat digunakan diantaranya *Algoritma genetika* (GA) dan *Particle swarm optimization* (PSO) (Dwiasnati & Devianto, 2021). Pada penelitian ini, penulis menggunakan metode *Algoritma Genetika* dalam optimasi parameter SVM karena mampu menemukan solusi yang optimal sehingga dihasilkan parameter terbaik dan dapat meningkatkan *accuracy* hasil klasifikasi.

*Algoritma Genetika* menurut (Suprayogi & Mahmudy, 2014) merupakan algoritma yang memanfaatkan proses seleksi alam yang dikemukakan oleh Charles Darwin. Algoritma genetika digunakan untuk melakukan simulasi komputer untuk mendapatkan solusi terbaik berdasarkan kandidat solusi yang terlihat. Dari representasi tersebut akan terbentuk individu-individu acak, membentuk suatu generasi. Populasi individu yang dihasilkan selanjutnya akan dievaluasi menggunakan *fitness* untuk mendapatkan populasi yang optimal.

Keuntungan penggunaan algoritma genetika terlihat dari kemudahan implementasi dan kemampuannya untuk menemukan solusi yang optimal dan bisa diterima secara cepat untuk masalah-masalah berdimensi tinggi. Penelitian menggunakan optimasi SVM-GA sebelumnya sudah dilakukan oleh (Kusumaningrum, 2017), yang berjudul “Optimasi Parameter *Support Vector Machine* Menggunakan *Genetic Algorithm* untuk Klasifikasi *Microarray Data*”, menggunakan *grid search* SVM pada data *colon* menghasilkan performa klasifikasi akurasi sebesar 75% sedangkan hasil performa klasifikasi SVM-GA pada data *colon* menghasilkan akurasi 90%. Dengan menggunakan SVM-GA menghasilkan akurasi yang lebih optimal dibandingkan dengan metode SVM tanpa GA.

## **B. Perumusan Masalah**

Berdasarkan latar belakang di atas, maka dapat dirumuskan suatu permasalahan yaitu, bagaimana pengklasifikasian data *stroke* menggunakan SVM dan optimasi parameter terbaik SVM menggunakan algoritma genetika. Hasil klasifikasi dinilai berdasarkan hasil performa klasifikasi yaitu *accuracy*, *precision*, *recall* dan *f-1 score*.

## **C. Batasan Masalah**

Pada penelitian ini penulis mengambil batasan masalah untuk mencegah meluasnya permasalahan dalam penelitian ini. Batasan penelitian ini adalah :

1. Data *stroke* yang digunakan dalam penelitian ini bersumber dari *kaggle.com*.
2. Klasifikasi pada penelitian ini hanya menggunakan fitur-fitur yang ada pada dataset yaitu *gender, age, hypertension, heart disease, ever married, work type, residence type, avg glucose level, bmi, smoking status* dan *stroke*.
3. Fungsi kernel yang digunakan pada SVM yaitu kernel *Radial Basic Function* (RBF).
4. Pengoptimasian menggunakan parameter C 1000 dan  $\gamma$  0.1.
5. Banyaknya populasi pada gen adalah 10.
6. Banyaknya *generations* yang digunakan adalah 10.

#### **D. Tujuan Penelitian**

Tujuan penelitian ini adalah mengklasifikasi penyakit *stroke* dan mengukur tingkat *accuracy, precision, recall* dan *f1-score* menggunakan metode *support vector machine* dan algoritma genetika sebagai pengoptimasian.

#### **E. Manfaat Penelitian**

Manfaat yang diharapkan dalam penelitian ini adalah menghasilkan model klasifikasi penyakit *stroke* menggunakan metode *support vector machine* dan algoritma genetika. Sehingga menjadi referensi bagi peneliti selanjutnya ketika akan melakukan penelitian dengan tema terkait.