

BAB I

PENDAHULUAN

1.1 Latar belakang

Klasifikasi adalah jenis analisis data yang membantu seseorang menentukan kelas atau label dari suatu sampel dalam klasifikasi [1]. Ini merupakan metode dari *supervised learning*, yang mencoba menemukan hubungan antara atribut masukan dan atribut target. Klasifikasi pertama kali diperkenalkan oleh Carolus Linnaeus untuk mengklasifikasikan spesies berdasarkan karakteristik fisik, dan kemudian ia dikenal sebagai bapak klasifikasi [2]. Klasifikasi memiliki tujuan untuk memprediksi kelas dari suatu objek yang belum diketahui sebelumnya [2], dan menurut [1], tujuannya adalah untuk meningkatkan kehandalan hasil dari data dengan mengelompokkan atau mengkategorikan data ke dalam kelas atau kategori tertentu berdasarkan karakteristik atau atribut yang dimiliki oleh data tersebut. Dalam klasifikasi, terdapat target variabel atau kategori. Metode-metode yang telah dikembangkan oleh peneliti untuk menyelesaikan kasus klasifikasi termasuk pohon keputusan, klasifikasi bayes atau yang dikenal sebagai *naive bayes*, serta metode lainnya seperti *support vector machine* (SVM) dan klasifikasi *K-nearest neighbor* (KNN). Dalam klasifikasi, terdapat tiga tahapan, yaitu pembangunan model, penerapan model, dan evaluasi. Proses klasifikasi melibatkan dua fase, yaitu fase training dan fase testing. Fase training digunakan untuk membangun model dan berfungsi sebagai data pembelajaran, sementara fase testing berfungsi untuk memvalidasi atau menguji model yang telah dibuat [3]. Salah satu penelitian yang terkait dengan klasifikasi dilakukan untuk mengklasifikasi penyakit Hipertensi oleh Hilda Zaqya Elnaz Putri. Penelitian tersebut bertujuan untuk mendeteksi dini apakah pasien terjangkit penyakit hipertensi atau tidak, dan menghasilkan akurasi yang sangat baik menggunakan metode SVM *grid search*, yaitu sebesar 89,42% [4].

Hipertensi menurut *World Health Organization* (WHO), merujuk pada kondisi tekanan darah dalam pembuluh darah mencapai 140/90 mmHg atau lebih tinggi [5]. Hipertensi dikenal sebagai "The Silent Killer", karena dapat menyerang tanpa menunjukkan gejala yang jelas, meningkatkan risiko terjadinya kondisi medis serius, termasuk penyakit jantung, gangguan otak, kerusakan ginjal, dan

berbagai penyakit lainnya [6]. Hipertensi menduduki peringkat sebagai penyebab utama penyebab kematian dini di seluruh dunia, memengaruhi lebih dari satu miliar orang, dengan satu dari empat pria dan satu dari lima wanita terkena dampaknya. Data Riskesdas tahun 2018 mencatat bahwa di Indonesia prevalensi hipertensi mencapai 34,11%, dengan estimasi kasus mencapai 63.309.620 orang dan angka kematian akibat hipertensi juga signifikan, mencapai 427.218 orang [7].

Diperlukan ketepatan dan akurasi yang baik dalam mengklasifikasi pasien hipertensi sebagai deteksi dini agar tidak terjadi kesalahan dalam memprediksi suatu penyakit, khususnya penyakit hipertensi. Para peneliti, terutama ahli kesehatan, menggunakan *data mining* sebagai alat untuk mencari solusi optimal dalam penanganan penyakit [8]. Penerapan algoritma yang didukung oleh berbagai metode bertujuan untuk meningkatkan akurasi prediksi. Sebagai contoh, penggunaan metode *Support Vector Machine* (SVM) pada klasifikasi penyakit hipertensi menghasilkan akurasi sebesar 89,22% menggunakan kernel RBF dan nilai *Area Under The Curve* (AUC) sebesar 89,42%. Hasil ini menunjukkan kinerja diagnostik yang sangat baik dibandingkan dengan penggunaan kernel linear yang menghasilkan akurasi sebesar 88,24% dan nilai *Area Under the Curve* (AUC) sebesar 88,30%. Dari perbandingan tersebut, dapat disimpulkan bahwa penggunaan kernel RBF memberikan kinerja yang lebih baik dalam klasifikasi penyakit hipertensi [9].

Dalam penelitian [10], metode *Support Vector Machine* (SVM) digunakan dalam diagnosis penyakit hepatitis. Hasil penelitian menunjukkan bahwa *Support Vector Machine* menggunakan kernel RBF menunjukkan akurasi yang lebih tinggi, yaitu sebesar 96%, sedangkan hasil akurasi menggunakan kernel linear juga cukup tinggi, yaitu sebesar 83%. Temuan ini menggambarkan keunggulan *Support Vector Machine* (SVM) dalam klasifikasi hepatitis, memberikan kontribusi penting terhadap pemahaman dan pengembangan diagnosis penyakit. Dengan tingkat akurasi yang tinggi, hasil ini menunjukkan potensi aplikatif *Support Vector Machine* (SVM) dalam mendukung praktik medis dan pengambilan keputusan terkait penanganan penyakit hepatitis.

Sebagai respons terhadap penelitian tersebut, penelitian ini bertujuan untuk melakukan klasifikasi penyakit Hipertensi. Metode yang akan digunakan adalah metode klasifikasi *Support Vector Machine* (SVM). Melalui penelitian ini, diharapkan hasil klasifikasi yang diperoleh dapat memberikan manfaat yang

signifikan bagi peneliti di masa yang akan datang.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Berdasarkan latar belakang di atas, menunjukkan bahwa metode *Support Vector Machine* (SVM) memiliki kemampuan yang baik dalam proses klasifikasi, khususnya pada kernel RBF. Oleh karena itu, penelitian ini akan menggunakan metode SVM untuk klasifikasi penyakit Hipertensi pada studi kasus UPT Puskesmas Modopuro, Kecamatan Mojosari, Kabupaten Mojokerto, Provinsi Jawa Timur dengan harapan menghasilkan akurasi yang maksimal.

1.2.2 Metode Usulan

Berdasarkan uraian masalah dan solusi diatas, solusi yang usulkan dalam penelitian ini adalah melakukan preprocessing data dengan menerapkan teknik encoding fitur serta menggunakan Standard Scaler untuk normalisasi. Setelah itu, data yang telah diproses akan diklasifikasikan menggunakan metode Support Vector Machine (SVM) dengan kernel RBF untuk meningkatkan akurasi dan efektivitas model.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang di atas, maka di dapatkan pertanyaan penelitian yaitu berapakah nilai akurasi dari penerapan *Support Vector Machine* (SVM) dalam klasifikasi penyakit hipertensi ?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Sumber data yang digunakan berasal dari UPT Puskesmas Modopuro, Kecamatan Mojosari, Kabupaten Mojokerto, Provinsi Jawa Timur, dengan rentang waktu lima bulan terakhir, yaitu mulai dari bulan Mei hingga September 2022.
2. Jumlah data keseluruhan yang digunakan adalah sebanyak 1749 data.
3. Pada penelitian ini menggunakan 9 atribut yaitu jenis kelamin, usia (dalam tahun), indeks massa tubuh (IMT), tinggi (cm), berat (Kg), tekanan sistolik (dalam mm Hg), tekanan diastolik (dalam mm Hg), tingkat nafas (per menit),

dan denyut nadi (per menit).

4. Klasifikasi pada penelitian ini akan dibagi menjadi dua kelas yaitu Penyakit Hipertensi terdapat 1000 data dan Bukan Penyakit Hipertensi terdapat 749 data.
5. Metode yang digunakan untuk klasifikasi adalah algoritma Support Vector Machine (SVM) dengan kernel RBF.

1.4 Tujuan Dan Manfaat

1.2.4 Tujuan

Berdasarkan pertanyaan dan uraian rumusan masalah diatas, dan tujuan penelitian ini bertujuan untuk mengetahui akurasi klasifikasi *Support Vector Machine* (SVM) menggunakan kernel *RBF*.

1.2.5 Manfaat

Manfaat dari penelitian ini agar dapat mengetahui nilai akurasi yang diperoleh dalam pembangunan sistem klasifikasi hipertensi menggunakan *Support Vector Machine* (SVM). Sehingga dapat dijadikan rujukan dalam pembuatan sistem diagnosa pasien Hipertensi.

1.2.6 Sistematika Laporan

BAB I PENDAHULUAN

Pada bab ini berisi dan menjelaskan tentang konteks latar belakang penelitian, perumusan masalah, tujuan dan manfaat penelitian, batasan penelitian, dan tata cara penyusunan dalam penulisan penelitian.

BAB II TINJAUAN PUSTAKA

Pada bab ini akan menguraikan landasan teori dan penelitian terdahulu yang terkait dengan penelitian yang diajukan penulis.

BAB III METODE USULAN

Pada bab ini akan menguraikan metode yang diusulkan mencakup arsitektur sistem, *dataset* yang digunakan, dan tahapan penelitian

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menjelaskan secara rinci mengenai penerapan dan hasil klasifikasi hipertensi yang didapatkan menggunakan metode *Support Vector Machine* (SVM).

BAB V PENUTUP

Bab ini berisikan kesimpulan yang telah dilakukan serta saran yang nantinya akan diharapkan dapat mengembangkan penelitian ini lebih lanjut.