

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

*The silent killer* atau juga biasa disebut pembunuh senyap adalah sebutan bagi penyakit tekanan darah tinggi atau hipertensi dalam istilah medis. Sebutan tersebut tentunya tidak hanya istilah semata, melainkan menggambarkan karakteristik penyakit yang dapat berkembang tanpa menunjukkan gejala yang signifikan hingga mencapai tahap berbahaya. Penderita hipertensi sering kali tidak menyadari kondisinya karena tidak merasakan sensasi khusus atau keluhan yang mengganggu aktivitas sehari-hari, padahal tekanan darah mereka telah melebihi batas normal dan sedang merusak organ vital secara perlahan namun pasti.

Salah satu penyakit kronis yang sering ditemukan di seluruh dunia yaitu adalah hipertensi. Jantung, stroke, dan gagal ginjal, adalah penyakit yang dapat menyebabkan kematian dan mengancam kesehatan masyarakat global [1]. *World Health Organization* (WHO) dengan jumlah 1,28 miliar orang dewasa dengan rentang umur 30-79 tahun di seluruh dunia menderita hipertensi, dengan dua pertiga di antaranya tinggal di negara berpendapatan rendah dan menengah [1]. Angka ini terus menunjukkan tren peningkatan yang signifikan seiring dengan perubahan gaya hidup modern yang cenderung tidak sehat, seperti konsumsi makanan tinggi garam dan lemak, kurangnya aktivitas fisik, stres berkepanjangan, serta kebiasaan merokok dan konsumsi alkohol berlebihan[2].

Di Indonesia, prevalensi hipertensi terus meningkat, terutama di kalangan populasi dewasa. Data dari Kementerian Kesehatan Indonesia menunjukkan bahwa prevalensi hipertensi di Indonesia mencapai 34,1% pada individu berusia  $\geq 18$  tahun [3]. Artinya, lebih dari sepertiga penduduk dewasa Indonesia berisiko mengalami komplikasi serius akibat tekanan darah tinggi. Dampak ekonomi dari hipertensi juga tidak dapat diremehkan, karena biaya pengobatan jangka panjang dan penanganan komplikasinya sangat membebani sistem kesehatan nasional serta keuangan keluarga penderita.

Penyakit hipertensi sering kali tidak terdeteksi pada tahap awal karena kurangnya gejala yang dirasakan oleh penderita, sehingga menyebabkan banyak kasus yang baru terdiagnosis setelah berkembang menjadi komplikasi yang lebih serius. Kondisi ini ternyata tidak didukung oleh masyarakat dengan cara pemeriksaan secara rutin dan minimnya akses ke layanan kesehatan berkualitas di beberapa daerah [2]. Ketika darah tinggi tidak terkontrol dalam jangka waktu yang lama, maka organ seperti jantung, otak, ginjal, mata, dan pembuluh darah mengalami kerusakan progresif yang dapat berujung pada kecacatan permanen bahkan kematian [3]. Oleh karena itu, deteksi dini sangat penting untuk mencegah dampak buruk yang lebih besar bagi kesehatan penderita.

Penggunaan teknologi informasi dan metode pembelajaran mesin dalam melakukan pengklasifikasian terhadap beberapa bidang telah banyak dilakukan. Teknologi tersebut yang nantinya dapat memungkinkan para pelaku medis dalam mendeteksi dan mendiagnosis hipertensi berdasarkan hasil rekam medis dari tenaga kesehatan.

*Support Vector Machine* (SVM) merupakan salah satu metode pembelajaran mesin yang populer dalam permasalahan klasifikasi, termasuk dalam klasifikasi data medis seperti hipertensi [4]. SVM bekerja dengan menemukan hyperplane yang optimal untuk memisahkan data. Keunggulan SVM terletak pada kemampuannya untuk menangani data berdimensi tinggi dan kompleks, serta kemampuannya dalam mengklasifikasikan data yang tidak seimbang [5]. Dibandingkan dengan metode pembelajaran mesin lainnya, SVM juga menunjukkan ketahanan yang baik terhadap overfitting [6], yang berarti model dapat digeneralisasi dengan baik untuk memprediksi masalah yang baru.

Pada tahun 2023, Mutmainnah dkk telah melakukan penelitian terkait klasifikasi dengan *Support Vector Machine* dan *Naive bayes* dengan ekstraksi fitur menggunakan *Fourier Descriptor* pada citra daun herbal. Proses klasifikasi dilakukan dengan membandingkan kedua metode tersebut, berdasarkan hasil klasifikasi yang diperoleh pada penelitian tersebut. Didapati bahwa metode SVM memberikan hasil performa yang sangat baik dengan akurasi sebesar 98% pada kernel *linier* [7]. Selain itu, Nur Aini dan Faisal juga melakukan penelitian terkait klasifikasi gangguan tidur dengan menggunakan SVM. Dalam penelitian tersebut,

peneliti menggunakan dataset *Kaggle Sleep Health and Lifestyle* yang terdiri dari 400 sampel data dan menerapkan SVM dengan tiga skenario pengujian serta empat jenis kernel berbeda (*Linear*, *Polynomial*, *RBF*, dan *Sigmoid*). Berdasarkan hasil klasifikasi yang diperoleh, metode SVM menunjukkan tingkat akurasi keseluruhan sebesar 92%. Lebih lanjut, didapati bahwa penggunaan "*Feature Importance*" dengan rasio data latih dan uji 80:20 menghasilkan akurasi tertinggi, yaitu 93% untuk semua jenis kernel yang diuji. Penelitian ini pun menyoroti bahwa SVM terbukti efektif dalam mengklasifikasikan data yang kompleks dan non-linear dibandingkan metode seperti *Naïve Bayes* yang digunakan pada studi sebelumnya.

Oleh karena itu, pemilihan *hyperparameter* pada model SVM adalah faktor yang sangat penting untuk mencapai performa klasifikasi yang optimal. Hasil klasifikasi yang buruk sering kali disebabkan oleh pengaturan parameter yang kurang tepat [8]. *Hyperparameter* merupakan parameter yang harus ditetapkan sebelum pelatihan model dan tidak dapat diubah selama pelatihan berlangsung [8]. Menentukan *hyperparameter* yang tepat bisa dilakukan melalui *manual tuning*, dengan catatan *manual tuning* diiringi dengan pemahaman yang mendalam tentang karakteristik parameter yang ideal untuk model dan dataset yang akan dilakukan permodelan. Penelitian [9], membandingkan dua teknik pemilihan parameter untuk *Support Vector Machines* (SVM) dalam prediksi *churn* pada layanan langganan. Dalam studi ini, kernel polinomial dipilih karena secara konsisten memberikan akurasi tertinggi dan kesalahan minimum terendah dalam prediksi *churn*, *appetency*, dan *up-selling* dibandingkan dengan kernel yang lain. Kemudian pada penelitian [10] juga melakukan perbandingan penggunaan kernel dan penentuan parameter pada empat dataset kesehatan. Pada penelitian tersebut penentuan parameter gamma didapatkan dengan menghitung  $\gamma = (1/\text{jumlah fitur})$ . Yang kemudian didapati kernel polinomial dengan *manual tuning* berhasil mendapatkan akurasi terbaik dibandingkan dengan penggunaan kernel lainnya.

Berdasarkan hasil perbandingan beberapa metode pada penelitian sebelumnya. Pada penelitian tugas akhir ini mengusulkan metode *Support Vector Machine* (SVM) sebagai metode untuk melakukan klasifikasi pada penyakit hipertensi. Penggunaan metode SVM untuk mengklasifikasi penyakit hipertensi dapat disimpulkan sangat efektif dalam penerapannya. Kemudian, pada penelitian tugas

akhir ini akan menggunakan Kernel *Polynomial*, kemudian untuk nilai parameter yang akan digunakan  $C = 10$ ,  $\gamma = 0.7$ ,  $d = 3$ .

## 1.2 Rumusan Masalah

Perumusan masalah yang dihasilkan berdasarkan latar belakang adalah seperti berikut:

### 1.2.1 Permasalahan

Tingginya prevalensi hipertensi di Indonesia menjadi faktor risiko utama penyakit kardiovaskular, stroke dan penyakit lainnya yang mengancam nyawa. Pendeteksian dini hipertensi sangat diperlukan untuk pencegahan komplikasi, namun banyak tenaga medis masih melakukan diagnosa dengan metode konvensional. Hal ini dapat menghambat identifikasi penderita hipertensi sejak dini dan mengurangi efektivitas intervensi preventif yang tepat waktu. Diperlukannya sistem untuk mendukung diagnosa hipertensi pada pasien.

### 1.2.2 Solusi Permasalahan

Penelitian ini mengusulkan penggunaan metode Support Vector Machine (SVM) untuk melakukan klasifikasi hipertensi berdasarkan parameter klinis. SVM dipilih sebagai pengklasifikasi utama karena kemampuannya dalam memisahkan data pasien hipertensi dan non-hipertensi dengan margin yang maksimal melalui penentuan hyperplane optimal. Untuk mengatasi tantangan yang dihadapi dalam klasifikasi data medis yang kompleks, penelitian ini akan mengimplementasikan algoritma SVM dengan fungsi kernel polinomial, dengan tuning manual hiperparameter agar mendapatkan hasil akurasi yang optimal pada model. Dengan penggunaan SVM, diharapkan klasifikasi status hipertensi menjadi lebih akurat dan dapat diandalkan dalam membantu pengambilan keputusan bagi tenaga medis dalam upaya deteksi dini hipertensi.

### 1.2.3 Pertanyaan Penelitian

1. Apakah metode *Support Vector Machine* dapat melakukan prediksi pada data penyakit hipertensi dengan baik ?
2. Bagaimana hasil akurasi algoritma *Support Vector Machine* yang didapati dari metrik pengujian *Confusion Matrix* ?

### 1.3 Tujuan dan Manfaat Penelitian

Tujuan dan manfaat yang dihasilkan dari penelitian tugas akhir ini adalah sebagai berikut.

#### 1.3.1 Tujuan Penelitian

Berikut adalah tujuan penelitian ini.

1. Mengetahui hasil akurasi metode SVM terhadap klasifikasi status penyakit hipertensi.

#### 1.3.2 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Membantu tenaga medis dan pihak terkait dalam mengidentifikasi pasien yang terduga mengidap hipertensi dengan lebih cepat & tepat, sehingga mencegah terjadinya risiko kesehatan yang lebih buruk dapat terjadi.
2. Dapat digunakan sebagai referensi atau rujukan untuk studi mendatang terkait aplikasi metode pembelajaran mesin dalam klasifikasi pada penyakit hipertensi.

### 1.4 Batasan Masalah

Dalam penelitian ini terdapat batasan-batasan yang diterapkan, antara lain:

1. Data yang digunakan pada penelitian tugas akhir ini adalah data penyakit hipertensi hasil survey dari program *The Behavioral Risk Factor Surveillance System* (BRFSS) oleh Pusat Pengendalian dan Pencegahan Penyakit (CDC) yang diambil dari platform penyedia data *Kaggle.com*.
2. Dataset hipertensi berjumlah 26083 data, dengan dua kelas yaitu 14274 sampel kelas hipertensi dan 11808 sampel kelas tidak hipertensi.
3. Total fitur yang digunakan sebanyak 13 variabel fitur dan 1 variabel target.
4. Metode yang digunakan dalam melakukan klasifikasi adalah *Support Vector Machine* (SVM).
5. Kernel yang digunakan adalah kernel *Polynomial*.
6. Parameter SVM yang digunakan meliputi  $C = 10$ ,  $\gamma = 0.7$ ,  $d = 3$ .

## **1.5 Sistematika Penulisan**

Sistematika penulisan yang digunakan pada proposal tugas akhir ini adalah sebagai berikut:

### **BAB I PENDAHULUAN**

Pada bab ini akan membahas tentang latar belakang dari masalah yang dibahas perumusan masalah yang terjadi hingga solusi dari permasalahan, tujuan dan manfaat penerapan atas solusi yang diajukan, batasan-batasan, dan penyusunan proposal yang sistematis.

### **BAB II KAJIAN PUSTAKA**

Pada bab ini akan menjelaskan teori yang dapat membantu penelitian ini dalam memecahkan permasalahan yang ada pada penelitian ini

### **BAB III METODE PENELITIAN**

Bab ini memberikan penjelasan terkait jenis dan sumber data, dan variabel serta tahapan analisis

### **BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini berisi pembahasan atau penjelasan yang lebih rinci mengenai implementasi model klasifikasi penyakit hipertensi menggunakan metode SVM yang telah diusulkan pada bab sebelumnya dan hasil pengujian dari model yang dihasilkan. Pengujian dilakukan untuk mengukur seberapa tepat model dalam menghasilkan *output* yang ditargetkan.

### **BAB V PENUTUP**

Bab ini menyajikan rangkuman atau kesimpulan dari implementasi yang telah dilakukan, serta saran rekomendasi agar penelitian ini bisa dikembangkan pada penelitian mendatang.