

BAB I

PENDAHULUAN

1.1. Latar Belakang

Banyaknya permasalahan kesehatan yang terjadi pada saat ini khususnya negara Indonesia disebabkan karena banyaknya masyarakat yang kurang peduli dengan kebersihan. Salah satu masalah kesehatan yang disebabkan oleh adanya gangguan kelenjar berupa organ kecil yang terletak di bagian depan leher serta mengakibatkan perubahan bentuk dan fungsi produksi hormon menjadi lebih berkurang atau terlalu banyak merupakan suatu penyakit yaitu tiroid. Penyakit ini berupa kelenjar endokrin yang dapat membahayakan kesehatan pada tubuh manusia hingga berujung kematian.

Tiroid adalah kelenjar yang berbentuk kupu-kupu di bagian bawah leher, atau juga dikenal sebagai kelenjar gondok. Kelenjar ini berfungsi sangat penting dalam proses metabolisme berbeda yang terjadi dalam tubuh [1]. Di Indonesia menunjukkan bahwa sebanyak 596 pasien mengalami gangguan tiroid di RSUP Salang tahun 2019 yang terdiri dari 484 (81,2%) pasien dengan hipertiroid dan 112 (18,8%) pasien dengan hipotiroid. Gangguan fungsi tiroid ini didominasi oleh rentang usia 41-50 tahun (29,9%), dengan hipertiroid lebih banyak terjadi pada usia muda yaitu 31-40 tahun (87,6%) dan 21-30 tahun (87,4%), sedangkan hipotiroid banyak terjadi pada kelompok usia tua yaitu 61-70 tahun (27,1%) dan 51-60 tahun (23,5%). Berdasarkan jenis kelamin, pasien gangguan fungsi tiroid lebih banyak pada perempuan 458 orang (76,8%) dibandingkan laki-laki 132 orang (23,2%) [2]. Sekitar 11% orang di India terkena penyakit ini yang merupakan jumlah yang sangat besar dibandingkan dengan 2% di Inggris dan 4,6% di Amerika Serikat. Hal ini disebabkan oleh fakta bahwa telah lama terjadi kekurangan yodium di antara orang-orang yang kurang mampu di negara tersebut [3].

Klasifikasi adalah suatu teknik menemukan suatu pola yang mampu memisahkan kelas data yang satu dengan yang lainnya untuk menentukan objek yang masuk dengan kategori tertentu dengan melihat kelakuan dan atribut data baru dengan hasil yang didapat untuk memberikan sejumlah aturan

[4]. Proses dari mencari suatu himpunan model yang dapat mendiskripsikan dan membedakan kelas-kelas data atau konsep tersebut untuk memprediksi kelas dari suatu objek dan mengetahui nilai akurasi [5].

Penelitian ini menggunakan teknik klasifikasi dengan membedakan kelas-kelas dataset yang sudah ada. Metode klasifikasi yang digunakan dalam penelitian ini adalah *metode support vector machine (SVM)* kelebihan dari metode SVM adalah metode ini dapat menghasilkan model klasifikasi yang baik, dalam meningkatkan nilai akurasi yang lebih baik dalam penelitian ini menggunakan Adaboost, kelebihan metode Adaboost itu sendiri adalah salah satu metode boosting yang mampu menyeimbangkan kelas dengan memberikan bobot pada tingkat eror yang dapat merubah distribusi data. Dan pada penelitian ini menggunakan Forward sequential select.

Penelitian yang dilakukan (Syarifah Jasmine Putri, Adetia Pratama, Kevin Attaqwa dan Rahmaddeni, 2023), perbandingan kinerja algoritma *naive bayes* dan *k-nearest neighbors (knn)* dalam memprediksi penyakit tiroid dengan menggunakan 3711 data dengan menggunakan *framework python* dalam mencari perbandingan sebuah *accuracy* pada kedua algoritma, menghasilkan algoritma *naive bayes* dengan data testing 63% dan data training 64% sedangkan akurasi yang dihasilkan pada lagoritma *k-nearest naighbors (knn)* adalah 63%. Pada penelitian (Agus Heri Yunial, 2020) yang berjudul “*Analisis Optimasi Algoritma Klasifikasi Support Vector Machine, Decision Trees, dan Neural Network Menggunakan Adaboost dan Bagging*” dengan data 45.211 dan didapat hasil optimasi adaboost meningkatkan nilai akurasi dari algoritma support vector machine dari 88.93% menjadi 89.10%, decision tree 90.55%, neural network 90.38% dan nilai akurasi dari bagging yaitu 88.93%.

Seiring dengan kemajuan teknologi dan perkembangan bidang kesehatan, pendekatan Machine Learning telah menjadi fokus utama dalam pengembangan model prediktif untuk diagnosis penyakit. Dalam konteks ini, penggunaan metode seleksi fitur telah menjadi penting untuk mengidentifikasi atribut atau fitur yang paling relevan dalam dataset klinis, sehingga meningkatkan akurasi dan interpretabilitas model.

Support Vector Machine (SVM) adalah salah satu algoritma Machine Learning yang telah terbukti efektif dalam klasifikasi data medis, termasuk dalam diagnosis penyakit tiroid. SVM bekerja dengan mencari batas keputusan yang optimal antara kelas-kelas yang berbeda dalam ruang fitur, sehingga memungkinkan untuk klasifikasi yang akurat.

Adaboost, di sisi lain, adalah metode ensemble learning yang memadukan beberapa model lemah menjadi satu model yang kuat. Dalam konteks prediksi penyakit tiroid, penggunaan SVM dengan Adaboost dapat meningkatkan kinerja model dengan memperhitungkan kekuatan dan kelemahan masing-masing model SVM yang terintegrasi.

Dengan memadukan teknik seleksi fitur dan penggunaan model Machine Learning seperti SVM dan SVM dengan Adaboost, diharapkan dapat dikembangkan sistem prediksi yang efisien dan akurat untuk diagnosis penyakit tiroid. Pengembangan model yang lebih baik dapat membantu praktisi medis dalam membuat keputusan yang lebih tepat dan memberikan perawatan yang lebih baik kepada pasien dengan penyakit tiroid. Oleh karena itu, penelitian tentang klasifikasi penyakit tiroid menggunakan seleksi fitur dan teknik Machine Learning, khususnya SVM dan SVM dengan Adaboost, menjadi topik yang menarik dan penting untuk diselidiki lebih lanjut.

Berdasarkan latar belakang dan penelitian yang dilakukan sebelumnya pada peneliti ini menggunakan metode *support vector machine (SVM)*, Adaboost dengan feature selection *Forward Sequential Select (FSS)*. Seleksi fitur itu sendiri dapat mempengaruhi tingkat akurasi prediksi karena dataset berisi sejumlah fitur, oleh karena itu dalam memprediksi penyakit tiroid dibutuhkan seleksi fitur sehingga akan didapat nilai akurasi yang lebih baik dan akurat dari penelitian sebelumnya.

1.2. Identifikasi Masalah

Tiroid adalah kelenjar yang berbentuk kupu-kupu di bagian bawah leher, atau juga dikenal sebagai kelenjar gondok.

- a. Belum banyaknya masyarakat yang mengetahui akan bahayanya penyakit tiroid.
- b. Kurangnya informasi terkait dengan penyakit tiroid.
- c. Menentukan apakah seseorang terkena penyakit tiroid dengan menghitung probabilitas dan nilai akurasi serta kemungkinan dari penyakit dan gejala-gejala yang timbul berdasarkan nilai yang diberikan oleh pakar.

1.3. Rumusan Masalah

Rumusan masalah berdasarkan identifikasi masalah dalam penelitian ini adalah Bagaimana mengetahui nilai akurasi yang optimal menggunakan Algoritma *Support Vector Machine (SVM)*, Adaboost dengan feature selection *Forward Sequential Select (FSS)* dalam memprediksi penyakit tiroid.

1.4. Batasan Masalah

Batasan masalah pada penelitian ini adalah :

- a. Algoritma klasifikasi yang digunakan adalah algoritma *Support Vector Machine (SVM)*, Adaboost dengan feature selection *Forward Sequential Select (FSS)* dengan tool *RapidMiner*. Data yang digunakan adalah dataset yang didapat dari situs [kaggle.com](https://www.kaggle.com)
- b. Hasil akhir yang didapat menentukan seseorang yang terkena penyakit tiroid dengan menghitung probabilitas yang ditimbulkan dari gejala-gejala yang timbul berdasarkan nilai yang diberikan oleh pakar.

1.5. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menghasilkan model algoritma yang memiliki nilai akurasi yang sangat tinggi, terkait dengan prediksi penyakit tiroid menggunakan Algoritma *Support Vector Machine (SVM)*, Adaboost dengan feature selection *Forward Sequential Select (FSS)* dalam memprediksi penyakit tiroid.

1.6. Manfaat Penelitian

Manfaat penelitian ini adalah :

- a. Bermanfaat untuk mengetahui seseorang yang terkena penyakit tiroid dengan menghitung probabilitas yang ditimbulkan dari gejala-gejala yang timbul berdasarkan nilai yang diberikan oleh pakar.
- b. Bagi institusi khususnya studi Magister Teknik Informatika dapat digunakan sebagai referensi ilmiah dalam penelitian selanjutnya dengan menerapkan teknik *Data Mining*.

1.7. Sistematika Penulisan

Proposal ini terdiri dari tiga bagian yang secara garis besar untuk sistematika penulisannya adalah sebagai berikut :

- a. Bab I Pendahuluan, yang berisi Latar Belakang Masalah, Identifikasi Masalah, Rumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian dan Sistematika Penulisan.
- b. Bab II Landasan Teori, yang berisi Penelitian Terkait dan Landasan Teori yang menjadi acuan dalam pelaksanaan penelitian.
- c. Bab III Metode Penelitian, yang berisi Metode atau Kerangka Kerja yang digunakan dalam penelitian ini. Serta, disajikan juga rencana penelitian berikut tahapan dan waktu pelaksanaannya.