

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penyakit jantung merupakan salah satu penyakit tidak menular yang mematikan. Dikutip dari data *Institute for Health Metrics and Evaluation (IHME)*, jumlah warga Indonesia yang meninggal dunia akibat penyakit jantung atau kardiovaskular mencapai angka 251,09 per 100.000 orang pada tahun 2019. Jumlah tersebut mengalami peningkatan sebanyak 1,25% dibandingkan tahun 2018 yaitu sebesar 247,99 per 100.000 penduduk. Menurut Wihastuti (2016) penyakit jantung merupakan penyakit yang terjadi saat arteri koroner mengalami masalah. Hal ini bisa disebabkan karena penumpukan plak dalam pembuluh darah arteri dan membuat dinding arteri menjadi keras dan tebal. Plak tersebut kemudian akan menyebabkan aliran darah ke tubuh menjadi terhambat atau berhenti. Jaringan yang bergantung pada arteri tersebut tidak mendapatkan oksigen serta nutrisi yang cukup.

Pada tahun 2022, BPJS kesehatan mencatat kasus penyakit kardiovaskular terbanyak di Indonesia adalah penyakit jantung dengan jumlah kasus sebesar 15,5 juta kasus. Kasus ini meningkat dibandingkan tahun sebelumnya. Dilihat dari data yang ada, tren kasus penyakit jantung di Indonesia terus mengalami peningkatan. Penyakit jantung merupakan penyakit penyebab kematian paling utama. Angka kematian akibat penyakit jantung diprediksi akan meningkat setiap tahunnya. Gaya hidup masyarakat Indonesia dan ketidaktahuan mengenai pendeteksian penyakit sejak dini menjadi salah satu faktor terjadinya peningkatan penyakit jantung tiap tahunnya (Hanifah dkk., 2021).

Upaya yang dapat dilakukan untuk mengatasi masalah ini salah satunya dengan melakukan diagnosis awal bagi masyarakat. Diagnosis awal penyakit jantung dapat menggunakan beberapa cara. Salah satunya dapat dilakukan menggunakan model prediktif. Model prediktif dapat digunakan untuk membantu melakukan diagnosis dini, prediksi risiko, dan rekomendasi perawatan yang tepat bagi individu yang berisiko. Model prediktif yang dapat digunakan untuk memprediksi penyakit dengan teknik klasifikasi antara lain adalah *Support Vector Machine, AdaBoost, K-Nearest Neighbor, Random Forest, Stochastic Gradient Descent, CN2 Rule*.

Selain beberapa metode di atas terdapat metode lain yaitu pohon keputusan dan *Naïve Bayes*. Pohon keputusan adalah algoritma pembelajaran mesin yang dapat menghasilkan struktur pohon keputusan. Konsep pohon keputusan adalah mengubah data menjadi pohon berhirarki dan aturan-aturan keputusan. Manfaat utama dalam pengambilan pohon keputusan adalah dapat mengambil keputusan yang rumit menjadi lebih simpel sehingga dapat dengan mudah menemukan solusi dari permasalahan yang ada (Nasrullah, 2021). Pada pohon keputusan, langkah awal klasifikasi ialah menetapkan pola dan fungsi yang dapat mendefinisikan dan mempartisi data satu dengan kelas data lainnya, pohon keputusan juga dapat digunakan untuk melakukan klasifikasi pada data yang belum memiliki kelas data tertentu (Sambani & Nuraeni, 2017).

Namun demikian, metode pohon keputusan mempunyai kelemahan, yaitu pada kelas dan kriteria dengan jumlah besar seringkali terjadi *overlap* sehingga membutuhkan waktu dan jumlah memori yang semakin besar. Untuk itu, metode pohon keputusan dapat dioptimalkan dengan *Particle swarm Optimization* (PSO). PSO merupakan algoritma pencarian berbasis populasi yang digunakan untuk menemukan solusi dengan melakukan pencarian terhadap solusi acak. PSO juga merupakan algoritma optimasi yang dapat membantu proses pengambilan keputusan. PSO adalah metode optimasi *heuristic* global yang diperkenalkan oleh Dr. Kennedy dan Eberhart pada tahun 1995 berdasarkan penelitian perilaku kawanan burung dan ikan (Tarigan dkk., 2019). Dengan PSO dapat mengoptimalkan *dataset* yang saling tumpang tindih dan dapat menyeleksi fitur atribut dengan jumlah besar (Saputra dkk., 2020).

Metode prediktif lain yang dapat digunakan adalah *Naïve Bayes*. *Naïve Bayes* adalah metode probabilistik yang memanfaatkan teorema *Bayes* untuk menghitung probabilitas penderita penyakit jantung berdasarkan gejala-gejala yang diamati. *Naïve Bayes* memprediksi kemungkinan kejadian di masa depan berdasarkan pengalaman di masa lalu. Metode ini dipilih karena mudah diterapkan dan bekerja mandiri, artinya satu fitur dalam data tidak tergantung pada fitur lain dalam data yang sama (Handoko & Neneng, 2021).

*Naïve Bayes* memiliki kelemahan pada atribut atau fitur independen yang sering salah, sehingga hasil estimasi probabilitas tidak dapat berjalan dengan

optimal. Salah satu cara yang dapat dilakukan agar hasilnya dapat optimal adalah dengan *Particle Swarm Optimization* (PSO). PSO digunakan untuk menyeleksi fitur (*feature selection*) atau atribut menggunakan bobot atribut yang telah dihitung dan atribut yang telah diseleksi akan diprediksi menggunakan algoritma *Naïve Bayes* (Prabowo & Mujono, 2018).

Penelitian ini dilakukan dengan dilatarbelakangi oleh kebutuhan untuk meningkatkan akurasi dan keandalan prediksi penyakit jantung. Dengan penggunaan PSO untuk mengoptimalkan parameter-parameter dalam pohon keputusan dan *Naïve Bayes*, diharapkan akan memungkinkan pengembangan model yang lebih akurat dalam memprediksi risiko penyakit jantung, serta memberikan wawasan yang lebih baik dalam faktor-faktor yang berkontribusi pada risiko tersebut. Penelitian sebelumnya terkait topik tersebut telah dilakukan oleh Sukron dkk (2021) dan Saputra dkk (2020).

Penelitian ini bertujuan untuk membandingkan keefektifan dari metode pohon keputusan dengan algoritma C4.5 dan *Naïve Bayes* berbasis PSO dalam memprediksi penyakit jantung. Fokus penelitian adalah pada perbandingan dua metode klasifikasi umum yang dapat digunakan untuk memprediksi penyakit jantung, yaitu pohon keputusan dengan algoritma C4.5 dan *Naïve Bayes*, dan akan dioptimalkan menggunakan algoritma *Particle Swarm Optimization* (PSO). Penelitian ini berbeda dengan penelitian sebelumnya yang dilakukan oleh Mutiara (2020) dengan judul “Algoritma Klasifikasi *Naive Bayes* Berbasis *Particle Swarm Optimization* Untuk Prediksi Penyakit *Tuberculosis* (Tb)” dan penelitian yang dilakukan oleh Ramdhani (2016) dengan judul “Penerapan *Particle Swarm Optimization* (PSO) Untuk Seleksi Atribut dalam Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis dengan Metode Algoritma C4.5”. Kedua penelitian tersebut hanya menggunakan salah satu metode dalam setiap penelitiannya dan menggunakan penyakit yang berbeda sedangkan pada penelitian ini membandingkan dua metode yaitu pohon keputusan dengan algoritma C4.5 dan *Naïve Bayes* berbasis *Particle Swarm Optimization* (PSO). Hasil penelitian ini dapat dimanfaatkan untuk pengembangan sistem perawatan kesehatan yang lebih efisien. Dengan memiliki model prediktif yang baik, perawatan pasien jantung

dapat dipersonalisasi berdasarkan risiko individu, sehingga mengurangi biaya perawatan yang tidak perlu dan meningkatkan hasil klinis.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah diuraikan, permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut.

1. Bagaimana memprediksi penyakit jantung menggunakan metode pohon keputusan dengan algoritma C4.5 berbasis PSO?
2. Bagaimana memprediksi penyakit jantung menggunakan metode *Naïve Bayes* berbasis PSO?
3. Bagaimana perbandingan hasil prediksi penyakit jantung dengan metode pohon keputusan dengan algoritma C4.5 dan *Naïve Bayes* berbasis PSO?

## 1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui cara kerja metode pohon keputusan dengan algoritma C4.5 dan *Naïve Bayes* berbasis PSO dalam memprediksi penyakit jantung dan membandingkan hasil prediksi dari keduanya.

## 1.4 Manfaat Penelitian

Penelitian ini bermanfaat untuk membantu memprediksi penyakit jantung sejak dini. Setelah mendapatkan hasil penelitian, nantinya metode yang lebih baik dalam memprediksi penyakit jantung dapat dikembangkan untuk bisa membuat web atau aplikasi yang mudah digunakan untuk pengguna.