

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Penyakit jantung koroner merupakan penyakit jantung yang disebabkan oleh penyempitan yang terjadi pada arteri koroner (Soeharto, 2001). Arteri yang mengalirkan darah ke otot jantung mengalami gangguan, sehingga jantung tidak mampu untuk memompa darah untuk memenuhi organ-organ vital. Salah satu penyebab terjadinya penyakit jantung koroner adalah kadar kolestrol yang tinggi.

Data *World Health Organization* (WHO) tahun 2015 menunjukkan lebih dari 17 juta orang di dunia meninggal akibat penyakit kardiovaskular atau 31% dari kematian di seluruh dunia. Dan di Indonesia angka kematian yang disebabkan oleh penyakit kardiovaskular yakni 7,4 juta (42,3%) diantaranya disebabkan oleh penyakit jantung koroner (PJK).

Data mining merupakan tahapan pengelolaan data yang menggunakan teknik statistika, kecerdasan buatan, *machine learning* untuk mengekstrak serta mengidentifikasi informasi yang bermanfaat beserta indikator penting dari berbagai *database* seperti *Kaggle*. Tujuan dilakukan *data mining* adalah untuk memahami lebih jauh mengenai data yang diamati.

Jaringan Saraf Tiruan (*Artificial Neural Networks*) diantaranya adalah salah satu dari banyak metode *data mining* yang merespon masukan dan mencoba meniru kinerja otak manusia. Dengan *data set* berupa variabel, Jaringan Saraf Tiruan akan memproses kemudian menghasilkan *output* tertentu. Yang menjadi sebuah keunggulan dari Jaringan Saraf Tiruan adalah mampu beradaptasi sehingga dapat belajar dari data masukan yang diberikan sehingga dapat memetakan hubungan antara *input* dan *output*-nya.

Extreme Learning Machine (ELM) adalah salah satu metode terbaru dari Jaringan Saraf Tiruan. ELM merupakan jaringan saraf tiruan *feedforward* dengan *single hidden layer* atau biasa disebut sebagai *single hidden layer feedforward neural networks* (SLFNs). Namun kekurangan pada jaringan saraf *feedforward neural network* (FFNN) adalah memakan waktu yang cukup banyak oleh karena itu metode ELM diciptakan guna untuk mengatasi kelemahan-kelemahan dari jaringan saraf tiruan *feedforward* dalam hal *learning speed*.

Algoritma C4.5 adalah algoritma yang pada umumnya digunakan guna untuk membentuk *decision tree*. *Decision tree* adalah salah satu metode klarifikasi yang paling populer karena mudah diinterpretasi oleh manusia. Algoritma C4.5 sendiri merupakan algoritma klasifikasi yang banyak digunakan karena memiliki kelebihan dibandingkan dengan algoritma yang lain. Kelebihan algoritma C4.5 adalah dapat menghasilkan pohon keputusan yang memiliki tingkat akurasi yang dapat diterima dan efisien dalam menangani atribut yang bertipe diskret atau numerik. Algoritma C4.5 memilih atribut sebagai *root* atau dikenal sebagai akar dan kemudian membuat cabang untuk setiap masing-masing nilai serta memilah tiap kasus ke dalam cabang dan proses akan dilakukan berulang kali agar tiap-tiap cabang sampai semua kasus pada cabang mempunyai kelas yang sama.

1.2 BATASAN MASALAH

Ruang lingkup pada penelitian ini adalah sebagai berikut:

- a. Data yang diperoleh berasal dari *Kaggle*
- b. Membentuk model aplikasi pada *Extreme Learning Machine* menggunakan *tool* MATLAB
- c. Menganalisis keakuratan hasil *Extreme Learning Machine* dengan menggunakan *Confusion Matrix*.

- d. Membandingkan hasil yang didapat dari *Extreme Learning Machine* dengan hasil metode algoritma C4.5

1.3 RUMUSAN MASALAH

Masalah yang ingin dipecahkan adalah sebagai berikut:

1. Apa saja kriteria yang dibutuhkan untuk mendeteksi penyakit jantung koroner?
2. Seberapa tinggi tingkat akurasi dari masing-masing metode?

1.4 TUJUAN PENELITIAN

Tujuan dilakukannya penelitian ini adalah sebagai berikut:

1. Mengimplementasikan *Extreme Learning Machine* agar dapat mengidentifikasi penyakit jantung koroner
2. Membandingkan keakuratan hasil prediksi *Extreme Learning Machine* dengan metode algoritma C4.5.

1.5 MANFAAT PENELITIAN

Manfaat dari dilakukan penelitian ini adalah membantu pengguna dalam mengetahui seberapa akurat metode algoritma C4.5 dan metode *Extreme Learning Machine* dalam mengidentifikasi penyakit jantung.

Dengan mengambil *data set* dari *Kaggle* maka akan diproses berdasarkan metode algoritma C4.5 dan metode *Extreme Learning Machine* untuk dibandingkan tingkat keakuratannya.

1.6 SISTEMATIKA PENULISAN

Tujuan sistematika penulisan adalah memberi kemudahan kepada pembaca untuk mengetahui isi dari tiap-tiap bagian penelitian. Bagian sistematika penulisan terdiri dari:

a) **BAB I PENDAHULUAN**

Pada bab ini dijelaskan beberapa hal pokok yang menjadi dasar dari penelitian atau dasar dirancangnya suatu penelitian, seperti: Latar Belakang, Masalah Pokok, Tujuan dan Manfaat, serta Batasan Masalah Penelitian.

b) **BAB II TINJAUAN PUSTAKA**

Pada bab ini dijelaskan dan disajikan kerangka teori yang mendukung analisis dan penelitian, yaitu rancangan penyelesaian masalah pokok yang akan dibahas.

c) **BAB III METODOLOGI PENELITIAN**

Pada bab ini dijelaskan beberapa hal pokok seperti: Kerangka Pikir, Pelaksanaan Kegiatan, Metode Pengumpulan Data, Metode Prediksi, Metode Diagnosis.

d) **BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini dijelaskan hasil pengolahan data atau hasil kegiatan akan penelitian yang dilakukan, serta pembahasan hasil penelitian.

e) **BAB V KESIMPULAN DAN SARAN**

Pada bab ini dijelaskan kesimpulan umum akan hasil penelitian yang disajikan dengan ringkas dan padat. Saran biasanya ditunjukkan kepada pihak yang memperoleh manfaat dari hasil penelitian, juga memuat saran tindak lanjut apa yang harus dilaksanakan, atau penelitian lanjutan yang masih perlu dilakukan.