

Sistem Cerdas untuk Mendeteksi Dini Penyakit Jantung Dengan Decision Tree

Slamet Kacung¹, Edi Prihartono²

^{1, 2}Teknik Informatika Universitas Dr. Soetomo Surabaya

¹slamet@unitomo.ac.id(*), ²edi.prihartono72@gmail.com

Abstract - Heart attack is the deadliest disease in the world including Indonesia. According to the report the heart Foundation Indonesia showed that the death toll reached more than 27 of 100 people due to heart disease. Early detection of heart disease is very needed considering the many people who suffer from heart disease on average already advanced stage. Intelligent system of early detection of heart disease is a method to know the symptoms that need to be alerted immediately so that heart disease could be known as early as possible. The methods used in this study using Decision Tree Classifier, the datasheet used are taken from the UCI Machine Learning Repository consisting of thirteen 270 instance, attribute input and 1 target attribute.

The results of this research will result in a decision tree that can help the community and or used as a reference for a doctor in diagnosing early heart disease. The second is this research can also predict a person can be diagnosed with heart disease or not by giving the input a few symptoms that are already established, the research results cannot replace an existing heart examination but at least it can help society in General nor the doctor.

Keyword: decision tree classifier, intelligent system

Abstrak - Penyakit jantung merupakan penyakit yang paling mematikan di dunia termasuk Indonesia. Menurut laporan Yayasan Jantung Indonesia menunjukkan bahwa angka kematian mencapai lebih dari 27 dari 100 orang disebabkan penyakit jantung. Deteksi dini penyakit jantung merupakan hal yang sangat diperlukan mengingat banyak sekali orang yang menderita penyakit jantung rata-rata sudah stadium lanjut. Sistem cerdas deteksi dini penyakit jantung merupakan metode untuk mengetahui gejala-gejala yang perlu diwaspadai dengan segera agar penyakit jantung dapat diketahui sedini mungkin. Metode yang digunakan dalam penelitian ini menggunakan Decision Tree Classifier, datasheet yang digunakan diambil dari UCI Mesin Learning Repository yang terdiri 270 instance, dari 13 attribute input dan 1 attribute target.

Hasil dari penelitian ini akan menghasilkan pohon keputusan yang dapat membantu masyarakat dan atau digunakan sebagai acuan seorang dokter dalam mendiagnosa awal penyakit jantung. Yang kedua adalah penelitian ini juga dapat memprediksi seseorang dapat didiagnosa penyakit jantung atau tidak dengan memberikan inputan beberapa gejala yang sudah ditetapkan, penelitian ini tidak dapat menggantikan hasil pemeriksaan jantung yang sudah ada tetapi paling tidak dapat membantu masyarakat secara umum atau pun dokter.

Kata Kunci : decision tree classifier, sistem cerdas

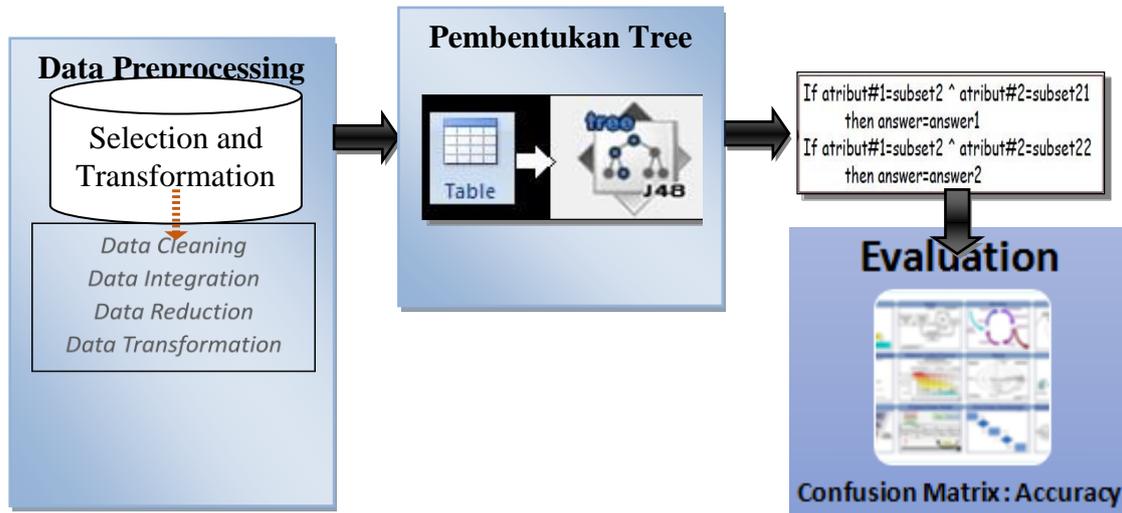
1. PENDAHULUAN

Penderita penyakit jantung di Indonesia semakin meningkat dari tahun ke tahun. Penyakit ini biasanya dimulai dengan rasa sakit yang tidak jelas, rasa tidak nyaman yang samar, atau rasa sesak dibagian tengah dada[1]. Jika ditangani lebih awal, penyakit ini lebih mudah diantisipasi. Akan tetapi pada sebagian besar kasus, penyakit diketahui dan ditangani saat sudah berada pada stadium lanjut. Hal ini disebabkan penyakit jantung sulit dideteksi. Jika sudah parah, maka Gangguan Jantung dapat dikenali dari Gejala Klinisnya saja, tetapi jika masih ringan dan baru cenderung untuk mengalami gangguan jantung, maka Treadmill Test yang dapat dilakukan di Laboratorium Klinik tertentu di bawah pengawasan Dokter Jantung merupakan salah satu cara untuk mendeteksi penyakit jantung. Pada Treadmill test penderita diberi beban latihan berjalan diatas alat treadmill lalu direkam dengan alat EKG pada saat yang bersamaan. Jadi test ini pada dasarnya merupakan gabungan antara latihan fisik dan monitor rekaman aktifitas listrik jantung selama latihan dengan akurasi antara 70-80% [2].

Oleh karena itu dalam penelitian ini kegiatan dokter dalam pengambilan keputusan untuk data diagnosa pasien jantung diekstrak, dan sistem cerdas dibangun berdasarkan pakar dan pengalaman manusia. Meskipun aplikasi ini dikeembangkan untuk mendiagnosa pasien jantung tapi keputsan utama pasien tersebut di diagnosa terkena jantung atau tidak adalah keputusan dokter. Aplikasi ini hanya membantu dokter untuk membuat penilaian alternatif atau sebagai alat bantu[3]. Algoritma decision tree sangat cocok untuk mengklasifikasikan terhadap pasien jantung dengan tingkat keparahan tertentu.

2. METODE

Dalam paper ini diajukan metode untuk mendeteksi dini penyakit jantung berdasarkan *Data Preprocessing*, Pembentukan *Decision Tree* kemudian di konfersi menjadi *Rule*, dan *evaluation*. Gambar 1. Menunjukkan langkah-langkah secara garis besar metode tersebut.



Gambar 1. Tahapan Metode Penelitian

Preprocessing

Dataset yang digunakan dalam penelitian ini diambil dari *UCI Mesin Learning Repository* merupakan public dataset yang sudah disepakati para peneliti data mining. Feature/attribute dari data tersebut di reduksi tujuannya agar menyederhanakan attribute dalam proses mining, misalnya attribute cp merupakan penyederhanaan dari attribute chest pain type yang memiliki empat type (*1 = Typical angina, 2 = Atypical*

angina, 3 = non-anginal pain, 4 = Asymptomatic), fbs penyederhanaan dari *Fasting blood sugar* yang memiliki nilai (1 jika fbs > 120 mg/dl dan 0 jika fbs < 120 mg/dl). Attribute sex memiliki nilai 1 = jika laki-laki, dan 0 = jika perempuan, sedangkan attribute tujuan (class) memiliki dua nilai yaitu (1 = Absence of heart disease (negative), 2 = Presence of heart disease (positif)). Untuk attribute lain dapat di lihat pada Tabel 1.

Tabel 1. Informasi Attribute[4].

No.	Attribute Selected	Attribute Name
1.	Age	Age
2.	Sex	Sex
3.	Cpt	chest pain type (4 values)
4.	Rbp	resting blood pressure
5.	Serchol	serum cholestoral in mg/dl
6.	Fb	fasting blood sugar > 120 mg/dl
7.	reselectres	electrocardiographic results (values 0,1,2)
8.	Mhra	maximum heart rate achieved
9.	Eia	exercise induced angina
10.	Oldpeak	oldpeak = ST depression induced by exercise relative to rest
11.	Slope	the slope of the peak exercise ST segment
12.	Nmv	number of major vessels (0-3) colored by flourosopy
13.	Thal	thal: 3 = normal; 6 = fixed defect; 7 = reversable defect
14.	Class	Class

Sumber : <http://www.ics.uci.edu/~mlern/MLRepository.html>

Pembentukan Tree

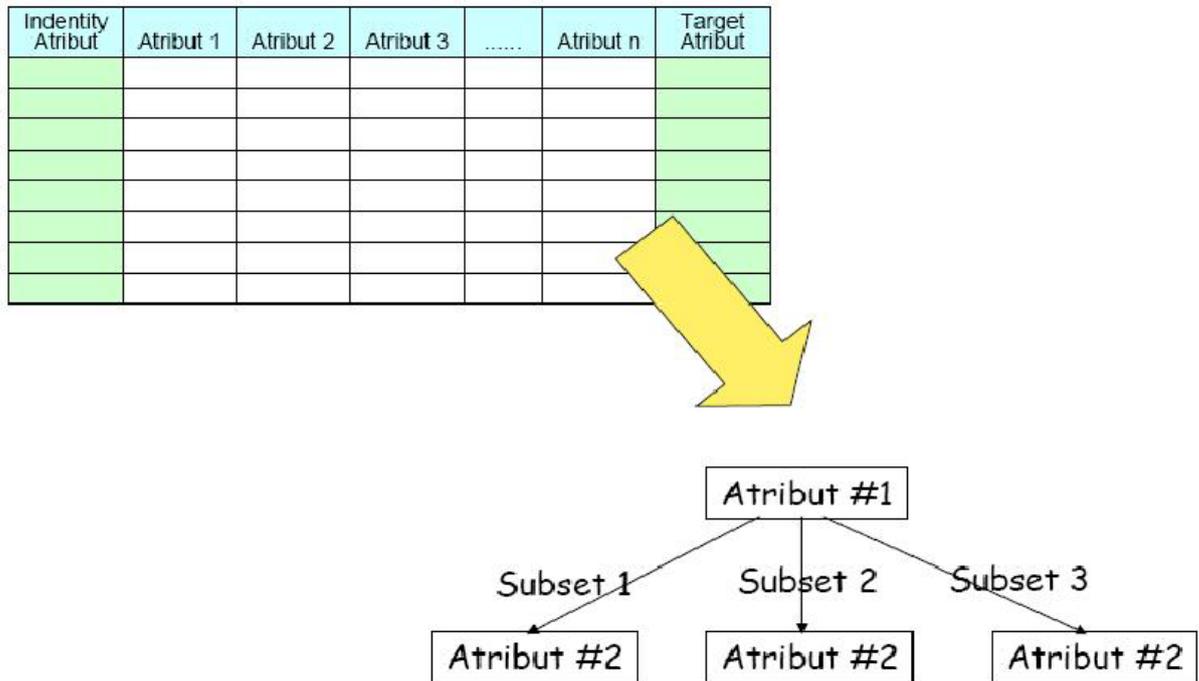
Proses pembentukan tree adalah mengubah bentuk data (tabel kasus) menjadi model pohon keputusan (*tree*), dan mengubah *tree* menjadi *rule* [5]. Secara umum untuk membangun *tree* adalah sebagai berikut :

- a) Memiih attribute sebagai root
Untuk memilih feature menjadi root menggunakan rumus :

$$Gain(S, A) = Entrophy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} x Entrophy(S_i)$$

$$Entrophy(S) = \sum_{i=1}^n -p_i x \log p_i$$

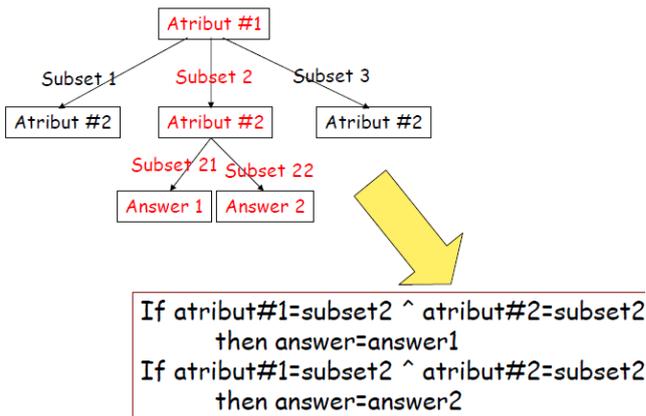
- b) Buat cabang untuk setiap nilai
- c) Bagi kasus dalam cabang
- d) Ulangi proses untuk setiap cabang, sehingga kasus pada setiap cabang mempunyai class yang sama.



Gambar 2. Proses Pembentukan Tree[5]

Pembentukan Rules

Proses pembentukan rule terjadi ketika pohon keputusan sudah terbentuk, tujuannya adalah untuk memudahkan hasil pemodelan sehingga hasil keputusan mudah dimengerti.



Gambar 3. Proses Pembentukan Rule[5]

Evaluasi

Evaluasi diperlukan untuk mengukur akurasi dan error dari sebuah model berdasarkan akurasi, kehandalan, dan kegunaan.

1. Akurasi
 - Mengukur seberapa baik model mengkorelasikan antara hasil dengan feature dalam data yang telah disediakan

- Ada beberapa model akurasi, tetapi semua model akurasi tergantung pada data yang digunakan.
2. Kehandalan :
 - Mengukur model data mining yang diterapkan pada dataset yang berbeda
 - Model data mining dapat diandalkan jika menghasilkan pola yang sama dari data testing yang disediakan
 3. Kegunaan

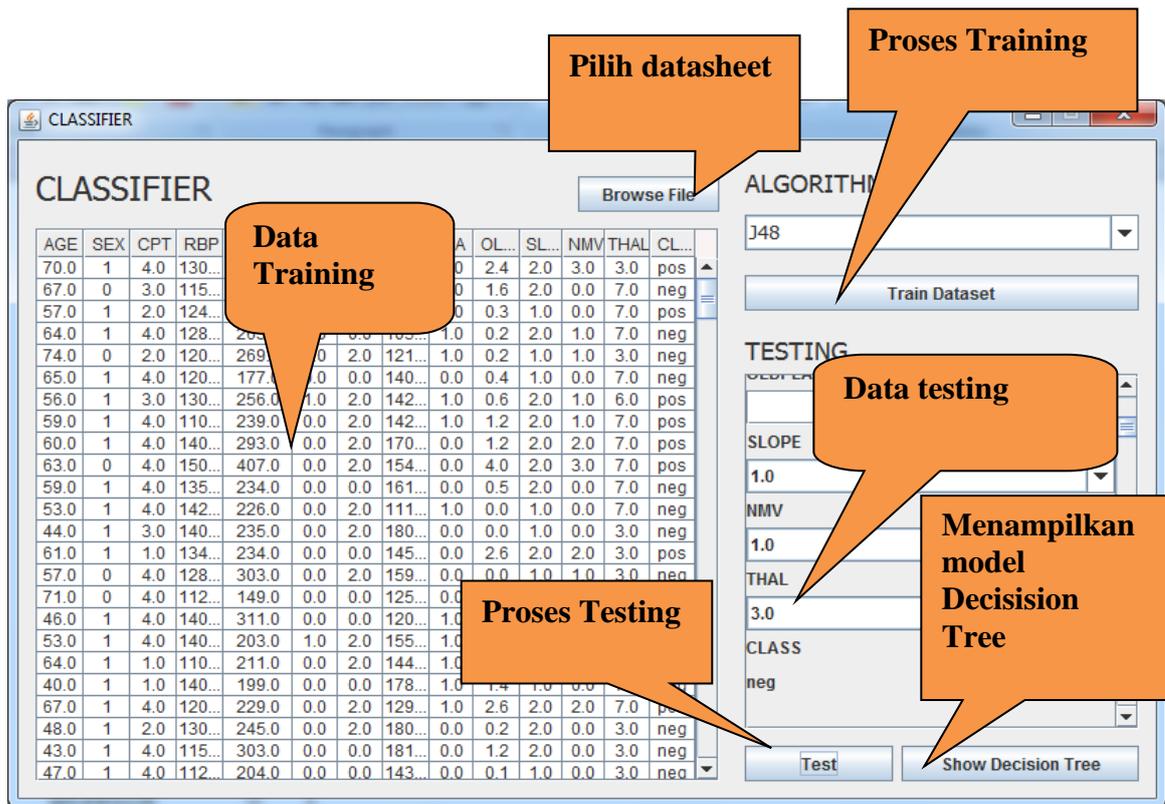
Kegunaan ini meliputi berbagai metric yang mengukur apakah model tersebut memberikan informasi yang berguna

Keseimbangan dari ketiganya sangat diperlukan karena belum tentu model yang akurat adalah handal, dan yang handal atau akurat belum tentu berguna.

III. HASIL DAN PEMBAHASAN

Hasil Pengetahuan

Pada eksperimen ini proses yang dilakukan adalah upload data yang akan dilakukan proses mining, kemudian proses training. Proses training ini merupakan proses yang harus dilakukan oleh mesin (komputer) agar mesin ini memiliki kemampuan untuk memprediksi. Pada proses training sejumlah data akan dibagi menjadi dua 80 % untuk training dan 20% untuk testing.



Gambar 4. Proses Training

IV. KESIMPULAN

- a) Aplikasi mampu membantu pihak sekolah / panitia PPDB dalam penerimaan dan pengelompokan kelas peserta didik baru di SMP Islam Parlaungan Waru Sidoarjo sesuai dengan kriteria yang telah ditentukan.
- b) Pengujian *White Box* dari perhitungan manual dengan perhitungan yang dilakukan oleh sistem yang menggunakan rumus metode *Multifactor Evaluation Proses* memberikan hasil penerimaan dan pengelompokan peserta didik yang sama dan sesuai dengan kriteria yang telah ditentukan, sehingga penerimaan dan pengelompokan peserta didik baru tidak perlu dilakukan secara manual dan dapat dilakukan dengan sistem pendukung keputusan penerimaan dan pengelompokan peserta didik.

- c) Aplikasi yang dibangun mampu menghasilkan laporan berupa hasil cetak / *print out* yang dapat digunakan sebagai bahan acuan dalam penerimaan dan pengelompokan peserta didik baru di SMP Islam Parlaungan Waru Sidoarjo.

DAFTAR PUSTAKA

- [1] Wikipedia. Retrieved November 16, 2015, available : https://id.wikipedia.org/wiki/Serangan_jantung
- [2] Deteksi Dini Penyakit Jantung dengan Treadmill Test, 2013, available : <http://rumahsakit.unair.ac.id/dokumen/Deteksi%20Dini%20Penyakit%20Jantung%20dengan%20Treadmill%20Test%201.pdf>
- [3] Visualization of Intelligent System using Decision Tree and Fuzzy Clustering for Heart Disease Early Detection, Information Systems International Conference (ISICO), 2013.
- [4] UCI Mechine Learning datasheet. Retrived December 1, 2015 available : <https://archive.ics.uci.edu/ml/datasets.html>
- [5] Basuki, Syarif, "Handout Kuliah Decission Tree" PENS, ITS. 2003