

ABSTRAK

Jauharotus Syafa'ah

(150411100022)

Dosen Pembimbing I : Sigit Susanto Putro, S.Kom., M.Kom.

Dosen Pembimbing II : Eka Mala Sari Rochman, S.Kom., M.Kom.

Jantung ialah salah satu organ yang bekerja pada sistem peredaran darah manusia. Penyakit atau kelainan jantung merupakan suatu keadaan dimana jantung tidak mampu bekerja dengan normal. Penyakit kardiovaskular adalah pelopor kematian nomor satu di dunia pertahun, yang ditimbulkan oleh disfungsi jantung dan pembuluh darah. Pada 2008, kiranya sebanyak 17,3 juta orang meninggal akibat penyakit kardiovaskular. Penyakit yang ditimbulkan akibat kegagalan fungsi jantung dan pembuluh darah disebut penyakit jantung. Penyakit tersebut dipengaruhi oleh beberapa faktor. Untuk mengidentifikasi apakah tubuh seseorang menderita penyakit jantung atau tidak, dapat dilakukan dengan berkunjung ke dokter untuk pemeriksaan kesehatan. Namun, tidak semua orang memiliki waktu, tenaga dan keberanian untuk memeriksakan keadaan tubuhnya, sehingga dilakukan sistem identifikasi untuk penyakit Jantung. Peneliti dapat mengetahui metode *Naive bayes classifier* sebagai metode yang tepat untuk membangun sebuah sistem identifikasi penyakit jantung , dan menjadikan penerapan metode *Naive Bayes Classifier* ini sebagai bahan literatur untuk penelitian-penelitian teknik klasifikasi selanjutnya. Pada sistem ini dapat mengklasifikasikan penderita penyakit jantung dan non jantung dari gejala klinis penyakit jantung, data yang digunakan berjumlah 1190. Observasi ini menggunakan perhitungan *Information gain* untuk proses pemilihan fitur, proses ini dipakai untuk memperkecil ukuran data dan meningkatkan hasil akurasi dari klasifikasi, Proses seleksi fitur jauh lebih bagus daripada tanpa menggunakan seleksi fitur. *K-fold cross validation* digunakan untuk memecah data membentuk data training dan data testing, penelitian ini menggunakan *kfold-5* dengan presentase 80% data training dan 20% data testing. Hasil akurasi rata-rata pengujian 5 *k-fold cross validation* diperoleh nilai akurasi sebesar 79.24%, presisi 78.77%, recall 80.74% dan f-measure 79.72%.

Kata kunci: *Jantung, Penyakit Jantung, Identifikasi, Klasifikasi, Information Gain, K-Fold Cross Validation, Naïve Bayes Classifier.*

ABSTRACT

Jauharotus Syafa'ah

(150411100022)

Advisor I : Sigit Susanto Putro, S.Kom., M.Kom.

Advisor II : Eka Mala Sari Rochman, S.Kom., M.Kom.

The heart is one of the organs that work in the human circulatory system. Heart disease or disorder is a condition in which the heart is unable to work normally. Cardiovascular disease is the number one cause of death in the world per year, caused by dysfunction of the heart and blood vessels. In 2008, an estimated 17.3 million people died from cardiovascular disease. Diseases caused by failure of the heart and blood vessels are called heart disease. The disease is influenced by several factors. To identify whether a person's body suffers from heart disease or not, it can be done by visiting a doctor for a health check. However, not everyone has the time, energy and courage to have their body checked, so an identification system for heart disease is carried out. Researchers can find out the Naive Bayes classifier method as the right method to build a heart disease identification system, and make the application of this Nave Bayes Classifier method as literature material for further classification technique studies. This system can classify patients with heart and non-cardiac diseases from clinical symptoms of heart disease, the data used is 1190. This observation uses Information gain calculations for the feature selection process, this process is used to reduce the size of the data and increase the accuracy of the classification results, the selection process feature is much nicer than without using feature selection. K-fold cross validation is used to break down data to form training data and testing data, this study uses kfold-5 with a percentage of 80% training data and 20% testing data. The results of the average accuracy of the 5 k-fold cross validation test obtained an accuracy value of 79.24%, precision 78.77%, recall 80.74% and f-measure 79.72%.

Keywords: Heart, Heart Disease, Identification, Classification, Information Gain, K-Fold Cross Validation, Naïve Bayes Classifier.