

ABSTRAK

Joshua Satria Jayanagara (01081170012)

PENGEMBANGAN SISTEM PAKAR MENGGUNAKAN METODE KLASIFIKASI DATA MINING PADA DATASET GEJALA PENYAKIT

(viii + 104 halaman: 45 gambar, 23 tabel, 2 lampiran)

Data mining merupakan sebuah proses menemukan sebuah pengetahuan berdasarkan data – data yang ada. *Data mining* sudah banyak digunakan dalam berbagai bidang, salah satunya adalah bidang medis. Sebagai contoh penerapannya, *data mining* dapat digunakan oleh seorang pasien yang ingin mengetahui penyakit yang mungkin dideritanya melalui data – data gejala yang ia alami. Beberapa penelitian sudah meneliti *data mining* untuk mendiagnosis satu penyakit, akan tetapi belum banyak penelitian yang meneliti beberapa penyakit sekaligus.

Penelitian ini bertujuan untuk merancang dan mengembangkan sebuah model sistem pakar yang dapat mendiagnosis penyakit berdasarkan gejala yang diberikan. Penelitian ini membandingkan tiga algoritma klasifikasi *data mining*, yaitu *Logistic Regression*, *Support Vector Machine*, dan *Naïve Bayes* untuk mencari model klasifikasi dengan nilai evaluasi paling baik. Model yang didapatkan akan digunakan dalam sistem pakar dalam mendiagnosis penyakit. Ketiga algoritma ini diuji untuk mendiagnosis enam penyakit, yang masing – masing penyakit memiliki gejala yang berbeda satu sama lain. Enam penyakit tersebut adalah penyakit obesitas, tiroid, kanker payudara, jantung, umum, dan COVID-19. Algoritma dengan performa terbaik dipilih untuk dipakai dalam sistem pakar membuat diagnosis.

Hasil akhir dari penelitian ini adalah sebuah sistem pakar yang dapat menerima input berupa gejala – gejala penyakit dari *user* dan menghasilkan diagnosis penyakit. Disisi lain, perbandingan dari performa algoritma yang diuji coba dapat menjadi masukan bagi penelitian selanjutnya dalam membuat model klasifikasi menggunakan algoritma terkait.

Kata kunci: *data mining*, klasifikasi, *Support Vector Machine*, *Logistic Regression*, *Naïve Bayes*, gejala penyakit, sistem pakar

Referensi: 42 (1985-2020)

ABSTRACT

Joshua Satria Jayanagara (01081170012)

EXPERT SYSTEM DEVELOPMENT USING CLASSIFICATION DATA MINING METHODS FOR DISEASE SYMPTOMS DATASET

(viii + 104 pages: 45 figures, 23 tables, 2 appendices)

Data mining is a process of finding knowledge based on existing data. Data mining has been widely used in various fields, one of which is the medical field. As an example of its application, data mining can be used by a patient who wants to find out what illnesses he/she might be suffering through data on his symptoms. Several studies have examined data mining to diagnose one disease, but not many studies have examined several diseases at once.

This research aims to design and develop an expert system model that can diagnose diseases based on the symptoms given. Here, three data mining classification algorithms are compared, namely Logistic Regression, Support Vector Machine, and Naïve Bayes, to determine which one will produce the classification model with the best evaluation results. The produced model will be used in the expert system in diagnosing disease. These three algorithms are tested to diagnose six diseases with their own symptoms. The six diseases are obesity, thyroid, breast cancer, heart disease, general disease, and COVID-19. The best performing algorithm is selected for use in the expert system that makes the diagnosis.

The final result of this research is an expert system that can receive input in the form of symptoms from the user and produce a disease diagnosis. On the other hand, the comparison of the performance of the algorithm being tested can be an input for further research in making classification models using related algorithms.

Keywords: data mining, classification, Support Vector Machine, Logistic Regression, Naïve Bayes, disease symptoms, expert system

References: 42 (1985-2020)