

INTISARI

IMPLEMENTASI METODE *VECTOR SPACE MODEL* DALAM SISTEM DIAGNOSA PENYAKIT INFEKSI BERDASARKAN ILMU PATOFISIOLOGI

Oleh

Imam Budi Santoso
12/331426/PA/14680

Salah satu penyakit yang membahayakan dan berkembang dengan cepat adalah penyakit infeksi yang menyerang tubuh manusia. Penyakit infeksi ditangani dengan berbagai jenis tindakan dan pendekatan khusus dalam ilmu kesehatan, tergantung tanda dan gejala yang dialami oleh pasien. Salah satu cara untuk mendeteksi penyakit infeksi secara cepat dan tepat adalah dengan membangun teknologi aplikasi yang dapat mempermudah para praktisi di bidang kesehatan dalam mempelajari dan menentukan kemungkinan penyakit yang dialami oleh pasien sebagai tolak ukur dalam proses penanganannya sekaligus menyediakan informasi tentang penyakit berdasarkan ilmu patofisiologi. Pada penelitian ini, dibangun suatu sistem diagnosa penyakit infeksi berdasarkan tanda dan gejala dalam ilmu patofisiologi.

Metode komputasi yang digunakan dalam penelitian ini adalah *Vector Space Model (VSM)*. Metode tersebut merupakan salah satu bagian dalam *text mining* yang akurat. Metode *VSM* dapat melihat tingkat kesamaan atau kedekatan suatu kata dengan melakukan pembobotan. Dokumen gejala yang terdapat pada basis data dipandang sebagai sebuah vektor yang memiliki jarak dan arah. Tingkat kecocokan antara data tanda dan gejala yang dimasukkan oleh pengguna dan dokumen yang terdapat dalam basis data dihitung dengan rumus *cosine similarity*. Dalam implementasinya, penelitian ini menggunakan beberapa algoritma pendukung, yaitu algoritma *stemming* Nazief Adriani dan algoritma pembobotan *TF-IDF (Term Frequency-Inverse Document Frequency)*.

Pengujian sistem dilakukan dengan menggunakan data yang dikumpulkan dari praktisi kesehatan. Terdapat 100 data penyakit yang digunakan dalam pengujian ini yang berdasarkan pada tanda dan gejala penyakit. Hasil penelitian ini menunjukkan bahwa pengujian berdasarkan *term* memiliki akurasi 90% dengan nilai similaritas penyakit yang cocok dengan diagnosa praktisi kesehatan mencapai 0.823281111, sedangkan pengujian berdasarkan frasa memiliki akurasi 88% dengan nilai similaritas penyakit yang cocok dengan diagnosa praktisi kesehatan mencapai 0.776355556.

Kata kunci : Penyakit Infeksi, *Vector Space Model*, *Text Mining*, Patofisiologi

ABSTRACT

THE IMPLEMENTATION OF VECTOR SPACE MODEL FOR INFECTIOUS DISEASE DIAGNOSIS SYSTEM BASED ON PATHOPHYSIOLOGY SCIENCE

By

Imam Budi Santoso
12/331426/PA/14680

One of the dangerous diseases and expanding rapidly is an infectious disease that attacks the human body. Infectious diseases treated by various types of actions and specific approach in the health sciences, depend on the signs and symptoms experienced by the patient. One way to detect infectious diseases quickly and accurately is to build technology applications that can facilitate practitioners in the field of health in studying and determining the likelihood of disease experienced by patients as a reference in the handling process as well as provide information about disease based on pathophysiology science. In this study, it has built a system of infectious disease diagnosis which is based on the signs and symptoms in the science of pathophysiology.

The computational method which is used in this research is Vector Space Model (VSM). The method is one part inaccurate text mining. VSM methods can see the degree of similarity or proximity of a word by weighting. Documents symptoms which are contained in the database are seen as a vector that has the distance and direction. A match rate between the signs and symptoms of the data entered by the user and the documents contained in the database are calculated by the formula of cosine similarity. In the implementation, this study uses several supports of algorithms, namely stemming Nazief-Adriani algorithm and TF-IDF (Term Frequency-Inverse Document Frequency) weighting algorithm.

System testing was done by using data collected from health practitioners. There were 100 disease data used in these tests, which are based on disease signs and symptoms. The study result showed that the test based on term reached 90% accuracy which had a similarity value of 0.823281111 that matched with the diagnosis of the health practitioners, while the other test result was based on phrase had an accuracy of 88% which had a similar value that matched the diagnosis of health practitioners reached 0.776355556.

Keywords : Infectious Diseases, Vector Space Model, Text Mining, Pathophysiology