

INTISARI

Pencarian Dokumen Teks Tanaman Obat Berdasarkan Kemiripan Manfaatnya Menggunakan Algoritma *Winnower*

Oleh

Muryanto

10/309060/PPA/03410

Kebiasaan yang terjadi di masyarakat dalam memilih tanaman obat untuk pengobatan, secara umum dilakukan dengan cara melihat manfaat dari tanaman-tanaman tersebut, kemudian dicocokkan dengan gejala-gejala sakit yang dirasakan selanjutnya mana yang paling mirip dengan gejala tersebut maka tanaman itulah yang akan dipilih untuk pengobatan.

Sistem ini dibangun untuk mendukung pencarian data tanaman obat berdasarkan gejala-gejala sakit. Dokumen gejala sakit dan tanaman obat ditentukan fingerprintnya dengan menggunakan algoritma *Winnower*. Algoritma *Winnower* menggunakan prinsip *over lapping* dari k-gram, untuk mendapatkan nilai hash dari dokumen, kemudian menggunakan pergeseran *window* untuk memilih nilai hash minimum dari setiap window untuk mendapatkan *fingerprint* dari dokumen tersebut. Adapun untuk menentukan tingkat kemiripan antara dokumen gejala sakit dan dokumen manfaat tanaman obat, maka digunakan rumusan *Jaccard Coefficient*, dimana jumlah *fingerprint* yang sama kedua dokumen dibagi dengan jumlah total *fingerprint* dari kedua dokumen tersebut.

Dari hasil pengujian terhadap kinerja, maka sistem ini memiliki kinerja yang cukup baik, hal ini ditunjukkan dengan nilai rata-rata *precision* sebesar 93,60%, rata-rata *recall* sebesar 80,60% , dan rata-rata *accuracy* sebesar 97,18%.

Kata Kunci : tanaman obat, *winnower*, *fingerprint*, kemiripan, *Jaccard Coefficient*

ABSTRACT

Text Document Searching on Medicinal Plants Based on Benefit Similarity Use Winnower Algorithm

By

Muryanto

10/309060/PPA/03410

Habits that have occurred in the community in the selection of medicinal plants for treatment, generally carried out in order to see the benefits of these plants, then paired with symptoms of pain that is felt most is most similar to the symptoms are plants that will be selected for treatment.

The system is built to support a medicinal plant data search based on the symptoms, based on the principle of fingerprint similarity of these two documents. Fingerprint of document is determined using algorithms Winnower. Winnower algorithm uses the principle of overlapping of k - gram to obtain a hash value of the document, and then use the shift window to select a minimum hash value of each window to get a fingerprint of the document. As for determining the degree of similarity between pain symptoms document and benefits of medicinal plants document, then used a formula Jaccard Coefficient, which the number of same fingerprints are divided by the total number of fingerprints.

The results of the performance tests these systems have the performance is quite good, as shown by the value of the average of precision 93.60%, the average of recall 80.60%, and the average of accuracy amounted to 97.18%.

Key Word : medicinal plants, Winnower, fingerprint, similarity, Jaccard Coefficient