

IDENTIFIKASI PENYAKIT HEPATITIS DENGAN PENDEKATAN AGGLOMERATIVE HIERARCHICAL CLUSTERING

ABSTRAK

Penyakit *hepatitis* merupakan salah satu jenis penyakit hati. Walaupun hepatitis memiliki tipe yang bermacam-macam, tetapi gejala seseorang terkena penyakit hepatitis sangat mirip sehingga sangat sulit untuk menentukan bahwa orang tersebut terkena penyakit hepatitis tipe A, B atau C. Pada bidang teknik informatika, penelitian terkait hepatitis sudah pernah dilakukan sebelumnya oleh Estu Karunianingtyas. Pada penelitian tersebut hanya mendapatkan akurasi 51,11 % karena hanya menggunakan data gejala yang tidak pasti merujuk pada satu penyakit saja.

Pada penelitian ini menambahkan data laboratorium sehingga mempengaruhi hasil akurasi dimana data laboratorium dapat mendiagnosa dengan benar jenis penyakit hepatitis. Berdasarkan hal tersebut maka sistem cerdas dibuat untuk mengelompokkan pasien yang terinfeksi hepatitis A, B dan C sesuai dengan data gejala dan data laboratoriumnya. Tahap pada penelitian ini menggunakan *knowledge discovery in databases* sehingga tujuan penelitian dapat tercapai. Pengelompokkan ini menggunakan algoritma *agglomerative hierarchical clustering* dengan pengukuran kemiripan *single*, *average* dan *complete linkage*.

Ada empat pengujian yang dilakukan untuk menghitung akurasi, yaitu perhitungan data laboratorium yang mendapatkan hasil 100 %, data gabungan gejala dan laboratorium dengan hasil 82,72 %, data laboratorium dan data gejala yang sudah diproses dengan *principal component analysis* dengan hasil 80,90 % serta data laboratorium dan data gejala hasil dari *principal component analysis* dengan 100 %. Hasil yang baik didapatkan oleh normalisasi [0-1] dengan pengukuran kemiripan *complete linkage*. Data laboratorium yang digunakan untuk pengelompokkan sangat membantu karena hasil akurasi yang dihasilkan cukup baik.

ABSTRACT

Hepatitis is one of kind the liver diseases. There are many various types of hepatitis. However, the symptoms of hepatitis are very similar because of that it was very difficult to determine that the person was infected by hepatitis A, hepatitis B and hepatitis C. In informatics engineering's field, the research about hepatitis diseases had been done by Estu Karunianingtyas. In that research, it just got accuracy about 51,11 % because it only used symptoms data that were not only indicated one disease.

This research added laboratory data to influence the result of accuracy in which laboratory data can diagnose the hepatitis diseases correctly. Based on the previous explanation, intelligent system was made in order to cluster the patients who were infected by hepatitis A, hepatitis B, and hepatitis C that were appropriate with the laboratory data and the symptoms data. The step of this research used knowledge discovery in databases so that the purpose of this research can be achieved. The clustering's algorithm that were used were agglomerative hierarchical clustering with similarity measure of single, average and complete linkage.

There were four tests that were used to calculate the accuracy: calculation of laboratory data that got the result of 100 %, the data with a combination of symptoms and laboratory that got the result of 82.72 %, symptoms and laboratory data that had been processed by principal component analysis that got the result of 80,90 %, and the last was result of laboratory data from principal component analysis that combine with symptoms data was 100 %. The good results were obtained by normalization [0-1] and used the similarity measure of complete linkage. The laboratory data which were used to cluster were very useful because the accuracy result was good enough.