

ABSTRAK

Data yang dimiliki oleh suatu sektor akan semakin bertumpuk apabila tidak dimanfaatkan sebaik mungkin. Untuk mengolah data dibutuhkan suatu teknik yang mampu mengolah data dengan skala besar. Maka dari itu, data mining adalah solusinya. Data mining diyakini mampu mengolah data berskala besar secara cepat sehingga banyak sektor yang menggunakan teknik ini, tak terkecuali pada sector kesehatan. Adapun salah satu penerapan data mining pada sector kesehatan adalah mendiagnosis suatu penyakit pada pasien. Untuk mendiagnosis penyakit, maka dibutuhkan pengklasifikasian agar pasien mengetahui secara jelas apakah didiagnosis mengidap suatu penyakit atau tidak. Salah satu metode data mining yang mampu mengelompokkan data secara jelas adalah metode klasifikasi. Metode ini akan memproses model yang dapat membedakan kelas secara jelas pada suatu data. Metode klasifikasi terdiri dari beberapa algoritma, namun pada penelitian ini hanya berfokus pada perbandingan antara dua algoritma yaitu *K-Nearest Neighbor* (KNN) dan *Decision Tree C4.5*. Kedua algoritma tersebut akan dipadukan dengan suatu optimasi yang dinamakan *Particle Swarm Optimization* (PSO). Hasil yang diperoleh menunjukkan bahwa *Decision Tree C4.5* dengan PSO memiliki tingkat kinerja yang lebih baik dibandingkan KNN dengan PSO. Hasil *accuracy* pada algoritma *Decision Tree C4.5* dengan PSO untuk percobaan pertama diperoleh nilai sebesar 84.18%, sedangkan KNN dengan PSO hanya sebesar 79.79%. Pada percobaan kedua, nilai *accuracy* untuk algoritma *Decision Tree C4.5* dengan PSO adalah 89.11%, sedangkan KNN dengan PSO hanya 87.90%. Pada percobaan ketiga, nilai *accuracy* pada algoritma *Decision Tree C4.5* dengan PSO sebesar 91.26%, sedangkan KNN dengan PSO hanya sebesar 89.11%.

Kata Kunci: Data mining, K-Nearest Neighbor, Decision Tree C4.5, Particle Swarm Optimization

ABSTRACT

The data owned by a sector will accumulate if it is not used as well as possible. To process the data required a technique that is able to process data on a large scale. Therefore, data mining is the solution. Data mining is believed to be able to process large-scale data quickly so that many sectors use this technique, including the health sector. One of the applications of data mining in the health sector is diagnosing a disease in patients. To diagnose a disease, classification is needed so that patients know clearly whether they are diagnosed with a disease or not. One method of data mining that is able to classify data clearly is the classification method. This method will process a model that can clearly distinguish classes in data. The classification method consists of several algorithms, but in this study only focuses on the comparison between the two algorithms, namely K-Nearest Neighbor (KNN) and Decision Tree C4.5. The two algorithms will be combined with an optimization called Particle Swarm Optimization (PSO). The results obtained indicate that Decision Tree C4.5 with PSO has a better level of performance than KNN with PSO. results Accuracy in the algorithm Decision Tree C4.5with PSO for the first experiment obtained a value of 84.18%, while KNN with PSO was only 79.79%. In the second experiment, the value accuracy for the algorithm Decision Tree C4.5with PSO was 89.11%, while KNN with PSO was only 87.90%. In the third experiment, the value accuracy in the algorithm Decision Tree C4.5with PSO was 91.26%, while KNN with PSO was only 89.11%.

Keywords: Data mining, K-Nearest Neighbor, Decision Tree C4.5, Particle Swarm Optimization