

## ABSTRAK

Pengolahan bahan pustaka dilakukan untuk mengelompokkan buku perpustakaan berdasarkan kriteria tertentu agar pencarian informasi mudah dilakukan. Salah satu jenisnya yaitu klasifikasi menggunakan *Dewey Decimal Classification* (DDC). DDC mengklasifikasikan bahan pustaka berdasarkan disiplin ilmunya. DDC membagi ilmu pengetahuan dengan sistem persepuluhan terdiri dari 10 kelas utama, 10 kelas divisi dan 10 kelas sub divisi. Berdasarkan hal tersebut, DDC memiliki lebih dari 900 kelas disiplin ilmu. Apabila dilakukan secara manual, klasifikasi bahan pustaka ke dalam kelas DDC akan sulit dilakukan. Oleh karena itu penelitian ini bertujuan untuk memudahkan klasifikasi buku ke dalam kelas DDC dengan menerapkan *text mining* dalam mencari persamaan pola judul bahan pustaka terhadap kelas DDC.

Untuk mendapatkan inti judul bahan pustaka, maka dilakukan *text preprocessing* dengan tahapan *case folding*, *tokenizing*, *filtering* dan *stemming*. Proses *stemming* dilakukan dengan algoritma Sastrawi. Kemudian hitung bobot dari bahan pustaka menggunakan algoritma pembobotan tf-idf. Hasil dari pembobotan akan digunakan dalam proses *clustering* dengan metode K-Medoids. Hasil clustering berupa data bahan pustaka yang telah dibagi dalam tiga kelompok tertentu. Selanjutnya proses klasifikasi data bahan pustaka menggunakan metode *Naive Bayes Classifier*, pada proses ini kelompok tertentu hasil *clustering* akan digunakan. Proses klasifikasi menghasilkan nomor kelas DDC yang dipilih berdasarkan nilai probabilitas tertinggi.

Hasil dari penelitian yang dilakukan yaitu pengolahan bahan pustaka kedalam kelas DDC dapat cepat dilakukan. Proses *Clustering* data latih dengan metode K-Medoids dapat mempercepat proses klasifikasi. Berdasarkan hasil pengujian yang dilakukan terhadap 300 data latih dan 75 data uji, metode gabungan *Naive Bayes Classifier* dan K-Medoids dapat diterapkan sebagai metode klasifikasi bahan pustaka ke dalam kelas DDC dengan akurasi 77,33%, presisi 81,81% dan *recall* 75,12% dengan waktu eksekusi rata-rata  $2,19 \times 10^{-7}$  detik.

**Kata Kunci :** ddc, k-medoids, *naive bayes classifier*

## ABSTRACT

*Processing of library materials is carried out to classify library books based on certain criteria so that finding information is easy. One type is classification using the Dewey Decimal Classification (DDC). DDC classifies library materials based on their disciplines. DDC divides knowledge with a tithing system consisting of 10 main classes, 10 division classes and 10 sub division classes. Based on this, DDC has more than 900 discipline classes. If done manually, the classification of library materials into the DDC class will be difficult. Therefore, this study aims to facilitate the classification of books into the DDC class by applying text mining in finding similarities in the pattern of library material titles to the DDC class.*

*To get the essence of the title of the library material, carried out text preprocessing is with the stages of case folding, tokenizing, filtering and stemming. The stemming process is carried out with the Sastrawi algorithm. Then calculate the weight of the library materials using the tf-idf weighting algorithm. The results of the weighting will be used in the process clustering with the K-Medoids method. The results of clustering in the form of library material data that have been divided into three specific groups. Furthermore, the process of classifying library material data uses Naive Bayes Classifier method, in this process certain groups of results clustering will be used. The classification process produces a DDC class number that is selected based on the highest probability value.*

*The results of the research conducted, namely the processing of library materials into the DDC class can be done quickly. The process of Clustering training data with the K-Medoids method can speed up the classification process. Based on the results of tests conducted on 300 training data and 75 test data, the combined method of Naive Bayes Classifier and K-Medoids can be applied as a method of classifying library materials into DDC class with an accuracy of 77.33%, precision 81.81% and recall 75, 12% with an average execution time of  $2.19 \times 10^{-7}$  seconds.*

**Keywords :** *ddc, k-medoids, nave bayes classifier*