

## ABSTRAK

Pengelolaan koleksi buku pada perpustakaan modern memerlukan sistem yang efisien untuk memudahkan pencarian dan pengelompokan buku. Perpustakaan Universitas Pembangunan Nasional Veteran Jakarta menghadapi kendala dalam mengorganisir ribuan judul buku dengan berbagai kategori. Penelitian ini bertujuan untuk melakukan klasterisasi judul buku menggunakan algoritma K-Means serta menerapkan teknik text mining. Data yang digunakan berupa 5715 data peminjaman buku. Proses penelitian mencakup tahap pra-pemrosesan seperti *case folding*, *tokenization*, normalisasi, *stopword removal* dan *stemming*. Kemudian, penelitian dilanjutkan dengan ekstraksi fitur menggunakan TF-IDF untuk mengubah data teks menjadi bentuk numerik.. Proses klasterisasi dilakukan dengan K-Means dan dievaluasi menggunakan Elbow Method serta Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal adalah  $k = 5$  dengan nilai DBI sebesar 2.103. Metode K-Means mampu mengelompokkan judul buku berdasarkan kemiripan topik secara efektif. Hasil ini diharapkan dapat membantu pengelolaan koleksi perpustakaan serta memudahkan pengguna dalam menemukan buku yang relevan.

**Kata kunci:** *text mining, clustering, K-Means, TF-IDF, Davies-Bouldin Index*

## ABSTRACT

*Management of book collections in modern libraries requires an efficient system to facilitate searching and grouping processes. Library of Universitas Pembangunan Nasional Veteran Jakarta faces challenges in organizing thousands of book titles across various categories. This study aims to cluster book titles using the K-Means algorithm and apply text mining techniques. The data used consists of 5,715 book borrowing records. The research process includes preprocessing stages such as case folding, tokenization, normalization, stopword removal, and stemming, followed by feature extraction using TF-IDF to convert text data into numerical form. Clustering process is carried out using K-Means and evaluated using the Elbow Method and Davies-Bouldin Index (DBI). Results show that the optimal number of clusters is  $k = 5$  with a DBI value of 2.103. K-Means method is able to effectively group book titles based on topic similarity. These results are expected to assist in managing library collections and help users find relevant books more easily.*

*Keywords: text mining, clustering, K-Means, TF-IDF, Davies-Bouldin Index*

