

ABSTRACT

Latent Dirichlet Allocation (LDA) in topic modeling aims to automatically and systematically discover hidden topics from text data. User reviews are unstructured text data, so a method is needed that can identify patterns and relationships between hidden topics in the data. LDA was chosen because it is a relatively simple probabilistic method, yet effective in topic modeling. The data used in this study were user reviews of the Halodoc application, the most widely used digital health service application in Indonesia, obtained through web scraping from the Google Play Store. Topic modeling was performed using the LDA method, with topic quality evaluation based on the Normalized Pointwise Mutual Information (NPMI) metric to determine the optimal number of topics. The results of the study show that there are four optimal topics, with a coherence score of 0.0167. This method is useful because it can simplify a large and unstructured collection of text data into concise, systematic, and easily interpretable information.

Keywords: *Topic Modeling, Latent Dirichlet Allocation, Topic Coherence, User Reviews, Halodoc*