I. Introduction

In the digital age, there has been a notable increase in the demand for online media. This reflects of how society is increasingly relying on the internet for information. Users frequently use online media since it is convenient and lets them quickly get current information. Consequently, a system that can find and suggest information, significantly news articles, pertinent and helpful to users is crucial. A sound news recommender system can improve the reading experience by guaranteeing that the retrieved content corresponds with their search.

Wang, Z. et al. [1] developed a news article topic recommender system using RAKE algorithm. Although RAKE algorithm used in previous studies successfully generated keywords searched by readers, the dataset used was very limited and only included political topics. As a result, the system needed help accommodating the diversity of topics of interest to readers. Previous studies have found that small datasets are more suitable for teaching purposes [2]. Therefore, a larger dataset is needed to study the recommender system's behavior. Previous research has used datasets from the BBC news portal using the Neural Network approach and the Collaborative Filtering paradigm. The study produced accuracy values for the baseline model based on estimates. The system built is compatible with the BBC service infrastructure and has succeeded in generating news recommender based on popularity [3]. In addition, Khotimah, N., and Girsang, A. S. [4] developed a recommender system for Indonesian News Articles using a genetic algorithm approach. The process selects important sentences from the text to create a summary based on the desired size. The results show that text summarization can be done well with a genetic algorithm. However, they have suggested developing an algorithm to summarize word selection in Indonesian documents. In previous studies, Latent Dirichlet Allocation (LDA) was used to extract topic information from news articles at different times. Experimental results have shown that this method is effective. However, this method is less suitable for managing the number of subjects that develop dynamically daily [5].

Until now, there have been many studies on developing news article recommender systems using various methods and algorithms such as RAKE, Neural Network, Genetic Algorithm, Collaborative Filtering Paradigm, and LDA. Although these studies have made progress in keyword selection, text summarization, and topic information extraction of news articles, there are still shortcomings regarding the dataset that is limited to specific topics, not multi-topic. To address this problem, we offer a news topic recommender system with a broader number of datasets covering several topics so that users can find information in the news quickly and efficiently. We examine the news topic recommender system over a certain period with the Indonesian News dataset, which is multi-topic. We use RAKE algorithm to find and extract the most important keywords from each article by utilizing Indonesian language documents and various topics covered in the dataset. Researchers have proven the algorithm to be highly efficient due to its flexibility in various language contexts and straightforward implementation. However, our study also has several limitations. Due to the unavailability of user interaction data, our approach only focuses on content-based recommendations by checking keyword similarities between articles and not implementing or evaluating a comprehensive end-to-end recommendation system. Additionally, these evaluation systems only focus on the similarity aspect of the extracted keywords, not on user-centric metrics, such as precision, recall, or user satisfaction.

This paper covers 5 sections. Chapter 1 explains general information about this study. Chapter 2 discusses the literature study on algorithms used in previous studies. Then, in Chapter 3, we explain the design and implementation of a news topic recommender system that utilizes RAKE for keyword extraction. Chapter 4, we present the experimental results, including system performance evaluation and comparison. Finally, in Chapter 5 we discuss the conclusion and suggestions for further research.