ABSTRAK

Pemilihan algoritma enkripsi yang tepat merupakan tantangan krusial dalam menjaga keamanan data, namun banyak pengguna menghadapi kesulitan karena keterbatasan pengetahuan teknis dan banyaknya literatur yang harus dianalisis. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi berbasis *Content-Based Filtering* (CBF) yang terintegrasi dengan teknik *text mining* untuk membantu proses pemilihan algoritma enkripsi secara lebih cepat, tepat, dan berbasis data ilmiah. Metodologi penelitian yang dilakukan meliputi pengumpulan 300 artikel ilmiah dari portal Garuda Kemdikbud melalui teknik *web scraping*, yang kemudian melalui tahapan *preprocessing* seperti *tokenisasi*, *stopword removal*, dan *case folding*. Data tersebut direpresentasikan menggunakan metode TF-IDF, sementara tingkat kemiripan antara kebutuhan pengguna dan literatur dihitung menggunakan *Cosine Similarity*. Hasil penelitian menunjukkan distribusi algoritma yang paling dominan dibahas dalam literatur adalah RSA (52 artikel), AES (40 artikel), dan RC4 (25 artikel). Sistem rekomendasi yang dibangun terbukti efektif karena mampu memberikan saran yang relevan sesuai kebutuhan spesifik, seperti merekomendasikan AES sebagai algoritma utama untuk kriteria "algoritma cepat untuk data sensitif". Pemanfaatan *text mining* dan CBF dalam sistem ini memberikan solusi yang lebih efektif dibandingkan metode manual karena sistem mampu melakukan analisis literatur secara mendalam dan otomatis, sehingga meminimalkan risiko kesalahan pemilihan yang dapat berdampak pada inefisiensi komputasi dan kerentanan keamanan data.

Kata Kunci: Algoritma Enkripsi, *Content-Based Filtering, Cosine Similarity*, Keamanan Data, *Text Mining*, TF-IDF

ABSTRACT

Selecting the right encryption algorithm is a crucial challenge in maintaining data security, yet many users face difficulties due to limited technical knowledge and the vast amount of literature required to analyze. This research aims to develop a recommendation system based on Content-Based Filtering (CBF) integrated with text mining techniques to facilitate faster, more accurate, and scientifically based encryption algorithm selection. The research methodology involved collecting 300 scientific articles from the Garuda portal of the Ministry of Education and Culture (Kemdikbud) using web scraping techniques, followed by preprocessing steps such as tokenization, stopword removal, and case-folding. The data was represented using the TF-IDF method, while the similarity between user requirements and the literature was calculated using Cosine Similarity. The results show that the most dominant algorithms discussed in the literature are RSA (52 articles), AES (40 articles), and RC4 (25 articles). The developed recommendation system proved effective because it was able to provide relevant suggestions tailored to specific needs, such as recommending AES as the primary algorithm for the "fast algorithm for sensitive data" criterion. The use of text mining and CBF in this system provides a more effective solution compared to manual methods because the system is able to perform literature analysis in depth and automatically, thereby minimizing the risk of selection which can have an impact on computational inefficiency and data security vulnerabilities.

Keywords*: Encryption Algorithm, Cosine Similarity, Content-Based Filtering, Data Security, Text Mining, TF-IDF,*