

ABSTRAK

Akses terhadap dokumen putusan pidana umum yang cepat dan akurat masih menjadi tantangan akibat kompleksitas struktur dokumen, keterbatasan pencarian berbasis kata kunci, serta variasi karakteristik panjang query yang digunakan oleh pengguna. Penelitian ini bertujuan untuk mengevaluasi efektivitas sistem temu kembali informasi dengan membandingkan metode *Term Frequency–Inverse Document Frequency* (TF-IDF), *Best Match 25* (BM25), dan *Hybrid TF-IDF* dan BM25 dalam pencarian dokumen putusan pidana umum. Dataset yang digunakan terdiri atas 749 dokumen putusan pidana umum Pengadilan Negeri di wilayah Madura yang diperoleh dari situs resmi Mahkamah Agung Republik Indonesia. Pengujian dilakukan melalui tiga skenario berdasarkan variasi panjang query, yaitu query pendek, query panjang menengah, dan query panjang, dengan evaluasi menggunakan metrik *precision*, *recall*, dan *F1-score* pada berbagai nilai Top-K. Hasil pengujian menunjukkan bahwa pada query pendek, ketiga metode menghasilkan kinerja yang setara dengan nilai *F1-score* tertinggi sebesar 0.69 pada Top-K 30. Pada query panjang menengah, metode BM25 menunjukkan kinerja terbaik dengan nilai *F1-score* tertinggi sebesar 0.66 pada Top-K 50, dibandingkan TF-IDF sebesar 0.62 dan *Hybrid TF-IDF* dan BM25 sebesar 0.65. Sementara itu, pada query panjang dan nilai Top-K yang lebih besar, *Hybrid TF-IDF* dan BM25 menunjukkan kinerja terbaik dengan nilai *F1-score* tertinggi sebesar 0.70 pada Top-K 50. Secara keseluruhan, hasil penelitian menunjukkan bahwa *Hybrid TF-IDF* dan BM25 tidak selalu lebih unggul dibandingkan metode tunggal, namun memberikan efektivitas yang lebih baik pada kondisi query panjang dengan kebutuhan cakupan hasil temu kembali yang lebih luas.

Kata Kunci : Temu Kembali Informasi, Dokumen Putusan Pidana Umum, TF-IDF, BM25

ABSTRACT

Access to general criminal court decision documents in a fast and accurate manner remains a challenge due to the complexity of document structures, limitations of keyword-based retrieval, and variations in user query length characteristics. This study aims to evaluate the effectiveness of an information retrieval system by comparing the Term Frequency–Inverse Document Frequency (TF-IDF), Best Match 25 (BM25), and a hybrid TF-IDF–BM25 approach in retrieving general criminal court decisions. The dataset consists of 749 general criminal court decision documents from District Courts in the Madura region, obtained from the official website of the Supreme Court of the Republic of Indonesia. The evaluation was conducted under three query length scenarios: short queries, medium-length queries, and long queries. System performance was measured using precision, recall, and F1-score across multiple Top-K values. The results indicate that for short queries, all three methods achieve comparable performance, with the highest F1-score of 0.69 at Top-K 30. For medium-length queries, BM25 demonstrates the best performance, achieving the highest F1-score of 0.66 at Top-K 50, compared to 0.62 for TF-IDF and 0.65 for the hybrid approach. In contrast, for long queries and larger Top-K values, the hybrid TF-IDF–BM25 method outperforms the others, achieving the highest F1-score of 0.70 at Top-K 50. Overall, the findings suggest that the hybrid TF-IDF–BM25 approach does not consistently outperform single methods but provides greater effectiveness in handling long queries where broader retrieval coverage is required.

Keywords: *Information Retrieval, Criminal Court Decision Documents, TF-IDF, BM25*