

ABSTRAK

Dokumen putusan perkara perdata agama di Indonesia memiliki struktur yang panjang dan kompleks, sehingga menyulitkan pembaca dalam menemukan informasi utama secara cepat dan efisien. Kondisi ini mendorong perlunya pemanfaatan teknologi *automatic text summarization* untuk menghasilkan ringkasan yang representatif dan informatif. Penelitian ini bertujuan untuk mengembangkan sistem peringkasan otomatis dokumen putusan perdata agama berbahasa Indonesia menggunakan pendekatan *extractive summarization*. Metode yang digunakan dalam penelitian ini adalah kombinasi IndoBERT sebagai *language representation* dengan arsitektur BERTSum untuk melakukan seleksi kalimat penting dalam dokumen. Data yang digunakan berupa dokumen putusan perkara perdata agama yang diperoleh dari Direktori Putusan Mahkamah Agung Republik Indonesia. Proses penelitian meliputi tahap *preprocessing*, *modeling*, dan evaluasi menggunakan metrik ROUGE yang mencakup ROUGE-1, ROUGE-2, dan ROUGE-L dengan fokus pada F1-score. Hasil penelitian menunjukkan bahwa model mampu menghasilkan ringkasan yang cukup representatif terhadap isi dokumen. Nilai rata-rata F1-score yang diperoleh sebesar 0,6915 untuk ROUGE-1, 0,5746 untuk ROUGE-2, dan 0,5792 untuk ROUGE-L menunjukkan bahwa model mampu menangkap informasi penting dengan tingkat kesesuaian yang cukup baik. Namun demikian, model masih memiliki keterbatasan dalam menjaga keterpaduan antar kalimat, terutama pada dokumen dengan panjang yang lebih besar. Secara keseluruhan, pendekatan *extractive summarization* berbasis IndoBERT dan BERTSum dapat menjadi solusi yang efektif untuk meningkatkan efisiensi dalam memahami dokumen hukum.

Kata kunci: *extractive summarization*, BERTSum, IndoBERT, dokumen hukum, ROUGE

ABSTRACT

Religious civil court decision documents in Indonesia generally have long and complex structures, making it difficult for readers to quickly and efficiently identify the main information. This condition highlights the need for automatic text summarization techniques to generate concise and representative summaries. This study aims to develop an automatic summarization system for Indonesian religious civil court decision documents using an extractive summarization approach. The proposed method combines IndoBERT as a language representation model with the BERTSum architecture for selecting important sentences from documents. The dataset consists of religious civil court decision documents obtained from the official Directory of Decisions of the Supreme Court of the Republic of Indonesia. The research process includes preprocessing, modeling, and evaluation using ROUGE metrics, namely ROUGE-1, ROUGE-2, and ROUGE-L, with a focus on F1-score. The results show that the model is able to generate summaries that are sufficiently representative of the document content. The average F1-scores obtained are 0.6915 for ROUGE-1, 0.5746 for ROUGE-2, and 0.5792 for ROUGE-L, indicating that the model can effectively capture important information with a reasonable level of similarity to the reference summaries. However, the model still faces limitations in maintaining sentence coherence, especially for longer and more complex documents. Overall, the extractive summarization approach based on IndoBERT and BERTSum can serve as an effective solution to improve efficiency in understanding legal documents.

Keywords: *extractive summarization, BERTSum, IndoBERT, legal documents, ROUGE*