

Temu-balik Dokumen Hukum dengan Model Neural Re-Ranker = Legal Document Retrieval with Neural Re-Ranker Model

Dimas Ichsanul Arifin, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920552652&lokasi=lokal>

Abstrak

Volume data hukum yang dihasilkan semakin besar setiap harinya. Oleh karena itu, kebutuhan akan sistem otomatis dan semi-otomatis, seperti sistem temu-balik informasi meningkat. Sistem temu-balik informasi dokumen hukum membantu praktisi hukum menemukan dokumen yang relevan dengan cepat dan efisien. Terkait hal tersebut, penelitian ini mengeksplorasi penggunaan model neural re-ranker pada sistem temu-balik informasi dokumen hukum dalam bahasa Inggris dan bahasa Indonesia. Tidak hanya itu, penelitian ini juga membahas beberapa pendekatan untuk meningkatkan efektivitas proses fine-tune dari model neural re-ranker. Model neural re-ranker dikembangkan untuk melakukan pengurutan ulang terhadap hasil pencarian awal yang didapat dari model pencocokan teks BM25. Implementasi ini menggunakan beberapa model neural re-ranker seperti BERT, IndoBERT, mBERT, dan XLM-RoBERTa yang melalui proses fine-tune. Hasil eksperimen menunjukkan bahwa model neural re-ranker BERT, IndoBERT, dan mBERT dapat meningkatkan performa dari sistem temu-balik informasi dokumen hukum yang sebelumnya hanya memanfaatkan model berbasis pencocokan teks seperti TF-IDF dan BM25. Peningkatan ini terlihat dari Skor Mean Average Percision (MAP) yang meningkat dari 0,760 menjadi 0,834 pada salah satu skenario yang dilakukan. Hal ini menunjukkan kinerja keseluruhan sistem temu-balik informasi yang lebih baik pada berbagai kueri. Sementara itu, pendekatan berupa pembekuan lapisan encoder berguna untuk meningkatkan efektifitas dari implementasi dari sistem temu-balik informasi yang memanfaatkan model neural re-ranker.The volume of legal data produced is getting bigger every day. Therefore, the need for automated and semi-automatic systems, such as information retrieval systems, is increasing. Legal document information retrieval systems help legal practitioners find relevant documents quickly and efficiently. In this regard, this research explores the use of the neural re-ranker model in a legal document information retrieval system in English and Indonesian. Not only that, this research also discusses several approaches to increase the effectiveness of the fine-tune process of the neural re-ranker model. A neural re-ranker model was developed to re-rank the initial search results obtained from the BM25 text matching model. This implementation uses several neural re-ranker models such as BERT, IndoBERT, mBERT, and XLM-RoBERTa which go through a fine-tune process. Experimental results show that the neural re-ranker models BERT, IndoBERT, and mBERT can improve the performance of legal document information retrieval systems that previously only utilized text matching-based models such as TF-IDF and BM25. This improvement can be seen from the Mean Average Percision (MAP) score which increased from 0,760 to 0,834 in one of the scenarios carried out. This indicates better overall performance of the information retrieval system on various queries. Meanwhile, the approach of freezing the encoder layer is useful for increasing the effectiveness of the implementation of an information retrieval system that utilizes the neural re-ranker model.