

ABSTRAK

Wijaya, Putu Agung Ananta Wijaya (2024), Retrieval Skripsi Menggunakan Modifikasi BM25 Berbasis Posisi Kata Adaptif.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I: Dr. I Made Gede Sunarya, S.Kom., M.Cs. dan Pembimbing II: Dr. I Gede Aris Gunadi, S.Si., M.Kom.

Kata-kata kunci: *BM25*, Beban Posisi Adaptif, Perpustakaan Digital, Sistem Rekomendasi

Pertumbuhan koleksi skripsi yang sangat cepat pada perpustakaan digital Universitas Udayana (e-Perpus UNUD), menyebabkan mekanisme pencarian berbasis kata kunci konvensional semakin tidak memadai dalam menemukan dokumen yang relevan. sistem pencarian repositori skripsi masih bertumpu pada pencocokan string sederhana yang mengabaikan konteks semantik dan struktur abstrak, sehingga menghasilkan banyak dokumen tidak relevan dan meningkatkan beban kognitif pengguna. Penelitian ini bertujuan mengembangkan dan mengevaluasi model BM25 Modifikasi berbasis posisi kata adaptif untuk meningkatkan akurasi sistem rekomendasi skripsi pada perpustakaan digital Universitas Udayana. Penelitian menggunakan desain eksperimen Information Retrieval dengan 50 kueri uji dan ground truth yang dianotasi secara manual terhadap korpus skripsi sarjana. Sebanyak 36 skenario pengujian dilakukan dengan memvariasikan metode temu kembali (TF-IDF, BM25, dan PoWeR-BM25), representasi data (judul, abstrak, dan judul-abstrak), serta teknik praproses teks (*stemming* dan tanpa *stemming*). Kinerja sistem dievaluasi menggunakan *Mean Average Precision* (MAP), *Precision*, *Recall*, dan *F1-Score*, kemudian model terbaik diintegrasikan ke e-Perpus UNUD melalui *Web Service API* berbasis *Django*. Hasil menunjukkan BM25 Modifikasi secara konsisten mengungguli BM25 standar dan TF-IDF, dengan MAP tertinggi 0,627 pada top_k = 50 menggunakan abstrak *stemming*. *Stemming* meningkatkan recall secara signifikan tanpa menurunkan presisi secara berarti, sementara penggunaan abstrak sebagai sumber data memberikan performa terbaik. Temuan ini menegaskan bahwa pembobotan posisi adaptif dalam abstrak meningkatkan efektivitas temu kembali dan menawarkan solusi praktis untuk sistem rekomendasi perpustakaan digital.

ABSTRACT

Wijaya, Putu Agung Ananta Wijaya (2024), Retrieval Skripsi Menggunakan Modifikasi BM25 Berbasis Posisi Kata Adaptif.

This thesis has been approved and examined by Supervisor I: Dr. I Made Gede Sunarya, S.Kom., M.Cs. dan Supervisor II: Dr. I Gede Aris Gunadi, S.Si., M.Kom.

Keywords: *BM25*, Adaptive Position Weighting, Digital Library, Recommendation System

The rapid growth of the undergraduate thesis collection at Udayana University's digital library (e-Perpus UNUD) has rendered conventional keyword-based search mechanisms increasingly inadequate for identifying relevant documents. Current thesis repository search systems still rely on simple string matching that ignores semantic context and abstract structure, resulting in numerous irrelevant documents and increasing the user's cognitive load. This research aims to develop and evaluate a Modified BM25 model based on adaptive word positions to improve the accuracy of thesis recommendation systems in the Udayana University digital library. The study employs an Information Retrieval experimental design using 50 test queries and manually annotated ground truth against a corpus of undergraduate theses. A total of 36 testing scenarios were conducted by varying retrieval methods (TF-IDF, BM25, and PoWeR-BM25), data representations (title, abstract, and title-abstract), and text preprocessing techniques (stemming and non-stemming). System performance was evaluated using Mean Average Precision (MAP), Precision, Recall, and F1-Score; subsequently, the best-performing model was integrated into e-Perpus UNUD via a Django-based Web Service API. Results indicate that the Modified BM25 consistently outperforms standard BM25 and TF-IDF, achieving the highest MAP of 0.627 at top_k = 50 using stemmed abstracts. Stemming significantly improves recall without substantially decreasing precision, while the use of abstracts as a data source yields the best performance. These findings confirm that adaptive position weighting within abstracts enhances retrieval effectiveness and offers a practical solution for digital library recommendation systems.