

## INTISARI

### ***QUERY EXPANSION MENGGUNAKAN THESAURUS DAN ASSOCIATION RULE PADA TEMU KEMBALI INFORMASI***

Oleh

Rifa'atus Shalihah

21/485543/PPA/06216

Sistem Temu Balik Informasi umumnya menggunakan query dalam pencarian dokumen. Jumlah dokumen yang dihasilkan sangat bergantung pada query yang dimasukkan oleh pengguna. Salah satu hal yang sering terjadi pada sistem pencarian dokumen teks adalah kosakata permintaan pengguna yang tidak cocok dengan dokumen yang menyebabkan tidak ada dokumen yang diambil. sistem temu balik informasi yang baik harus bisa menangani adanya kesenjangan kosa kata dan query yang ada antara dokumen dan query. Query Expansion adalah salah satu teknik dalam temu balik informasi yang bertujuan untuk meningkatkan hasil pencarian dengan memperluas atau memperkaya kueri pengguna.

Beberapa tahapan yang dilakukan pada penelitian query expansion ini adalah preprocessing untuk semua data dokumen dan query. Setelah itu perhitungan TF-IDF, Dokumen Vektor, dan proses untuk perhitungan similarity dengan metode cosine similarity. Query expansion yang digunakan pada penelitian ini yaitu dengan menambahkan sinonim dari query yang dimasukkan menggunakan sinonim dari thesaurus. Selain sinonim, query expansion lain yang akan digunakan adalah association rule fp-growth dengan melihat keterkaitan antar kata.

Pengujian dilakukan dengan 30 kata kunci berdasarkan bab yang ada pada data hadis yang akan dijadikan acuan. Dari hasil pengujian didapatkan nilai rata-rata *precision*, rata-rata *recall*, dan rata-rata *accuracy* sebesar 0.372, 0.4952 dan 0.9383 dengan temu balik informasi tanpa *query expansion*. Selanjutnya nilai rata-rata *precision* dan rata-rata *recall* dan rata-rata *accuracy* sebesar 0.2053, 0.6320 dan 0.8019 dengan temu balik informasi dengan *query expansion* sinonim *thesaurus*. Selanjutnya nilai rata-rata *precision*, rata-rata *recall* dan rata-rata *accuracy* sebesar 0.2472, 0.6545, dan 0.5861 dengan temu balik informasi dengan *query expansion association rule fp-growth*. Selanjutnya nilai rata-rata *precision*, rata-rata *recall* dan rata-rata *accuracy* sebesar 0.1513, 0.6688, dan 0.6255 dengan temu balik informasi dengan *query expansion* gabungan sinonim *thesaurus* dan *association rule fp-growth*.

**Kata Kunci:** *Query Expansion, Thesaurus, Association Rule , Precision and Recall, Temu Kembali Informasi*

## ABSTRACT

### QUERY EXPANSION USING THESAURUS AND ASSOCIATION RULES IN INFORMATION RETRIEVAL

By

Rifa'atus Shalihah

21/485543/PPA/06216

Information Retrieval Systems generally rely on queries for document searches. The number of documents retrieved highly depends on the query entered by the user. One common issue in text document retrieval systems is the mismatch between user query and document terms, resulting in zero retrieved documents. A proficient information retrieval system should effectively address vocabulary gaps and discrepancies between the query and documents. Query Expansion is a technique in information retrieval aimed at enhancing search results by broadening or enriching user queries.

The stages involved in this query expansion research include preprocessing for all document data and queries, followed by TF-IDF calculation, Document Vectorization, and the process of computing similarity using the cosine similarity method. Query expansion applied in this study involves augmenting synonyms from the thesaurus to the entered query. In addition to synonyms, another query expansion method utilized is the association rule fp-growth, considering the associations among words.

The testing was conducted using 30 keywords based on the chapters present in the hadith data, which served as the reference. The average precision, recall, and accuracy values obtained were 0.372, 0.4952, and 0.9383, respectively, for information retrieval without query expansion. Subsequently, the average precision, recall, and accuracy values were 0.2053, 0.6320, and 0.8019 with thesaurus synonym query expansion. Moreover, the average precision, recall, and accuracy values were 0.2472, 0.6545, and 0.5861, respectively, using the association rule fp-growth for query expansion. Lastly, the average precision, recall, and accuracy values stood at 0.1513, 0.6688, and 0.6255, respectively, for query expansion employing a combination of thesaurus synonyms and association rule fp-growth.

**Keywords: Query Expansion, Thesaurus, Association Rule, Precision and Recall, Accuracy, Information Retrieval**