Abstrak

Pada hakekatnya sebuah dokumen teks merupakan suatu jenis basis data yang tidak terstruktur, karena tidak memiliki field-field seperti halnya basis data konvensional. Berbedanya topik sebuah dokumen dengan dokumen yang lain bisa diartikan bahwa dokumen-dokumen tersebut menyimpan informasi yang berbeda.

Kategorisasi dokumen text merupakan upaya untuk mengelompokan dokumen kedalam kelompok-kelompok yang sudah terdefinisi. Saat ini sudah banyak metode kategorisasi/klasifikasi dokumen seperti *classifier K-Nearest Neighbor (KNN)*, *Bayesian Classifier, Decision Tree* dan metode *Rocchio* yang biasa digunakan utuk permasalahan klasifikasi.

Permasalahan kategorisasi secara hirarkis sebenarnya lebih banyak ditemui dalam dunia nyata, seperti penyimpanan file-file digital dalam folder-folder yang biasanya tersusun hirarkis. Salah satu implementasi kategorisasi ini adalah dengan memanfaatkan *Fuzzy Relational Thesaurus* (FRT) sebagai struktur hirarki kelas dalam kategorisasi.

Tugas akhir ini mengimplementasikan metode klasifikasi dokumen teks yang menggunakan FRT sebagai hirarki topiknya dan memanfaatkan metoda *Rocchio* sebagai metode bantuan pembentuk klassifiernya. Hasil pengujian menunjukan proses *training* untuk memilih fitur terbaik dalam metode FRT bisa menghasilkan *classifier* yang lebih baik dari *classifier* metode *Rocchio*.

Kata kunci : *kategorisasi teks hirarkis*, *Metode Rocchio*, *Fuzzy Relational Thesaurus.*