

## Abstrak

Saat ini jumlah informasi seperti artikel berita yang ada di dalam web terus berkembang dengan cepat. Banyaknya jumlah informasi tersebut bisa menyebabkan pengguna mengalami kesulitan dalam mencari informasi yang penting. Salah satu komponen *Data Mining* yaitu *Text Categorization* yang merupakan suatu cara untuk mengelompokkan dokumen ke dalam kelas-kelas tertentu. Salah satu penggunaan *Text Categorization* adalah *Klasifikasi Hirarki* yaitu suatu cara memberikan kelas hirarki atau kelas spesifik ke suatu dokumen sehingga dokumen dapat terorganisasi dengan baik.

Beberapa algoritma yang digunakan dalam klasifikasi hirarki adalah algoritma *top down with parent preference*, *bottom up with parent support*, *bottom up with threshold*. Dengan ketiga algoritma tersebut, dokumen akan dikelompokkan dengan memperhatikan hubungan tiap kelas yang ada dalam *tree hirarki* yang telah ditentukan sebelumnya.

Pada Tugas Akhir ini dilakukan perbandingan performansi antara ketiga algoritma dengan menggunakan beberapa metode klasifikasi sebagai pembentuk classifiernya, yaitu metode *Centroid Sum* dan *Naive Bayes Multinomial*. Parameter performansi yang digunakan adalah *hierarchical precision*, *hierarchical recall*, dan *f-measure*. Hasil percobaan menunjukkan bahwa algoritma *top down with parent preference* memberikan performansi yang lebih baik dibandingkan dengan algoritma *bottom up with threshold*, algoritma *top down with parent preference* dan *bottom up with parent support* mempunyai nilai performansi yang sama.

**Kata Kunci** : *text categorization*, *klasifikasi hirarki*, *top down with parent preference*, *bottom up with parent support*, *bottom up with threshold*, *tree hirarki*, *f-measure*.