ABSTRAKSI

Teknologi informasi yang berkembang dengan cepat memungkinkan bagi banyak instansi atau perusahaan untuk mengumpulkan dan menyimpan data dengan jumlah yang sangat besar. Data mining merupakan suatu proses untuk mengekstrak dan menganalisis data dalam jumlah besar ini, untuk mendapatkan pola-pola yang menarik sehingga didapat pengetahuan dari data-data tersebut. Salah satu fungsionalitas data mining adalah klasifikasi, dimana proses ini adalah untuk menemukan suatu gambaran model yang membedakan kelas-kelas data, agar dapat memprediksi objek-objek yang kelasnya belum diketahui.

Tugas Akhir ini mengimplementasikan algoritma *MIND* sebagai salah satu algoritma klasifikasi dengan *decision tree* untuk data dengan jumlah besar dan melakukan analisis terhadap performansi meliputi akurasi pengklasifikasian, waktu proses (waktu pengklasifikasian) dan skalabilitas, yang hasilnya dibandingkan dengan algoritma *SPRINT*. Implementasi ini sendiri dilakukan menggunakan *SQL Server 2000* dan *Delphi 6*. Data yang digunakan untuk pengujian beragam mulai dari jumlah *record* sebanyak 57 sampai 3.163, jumlah atribut 10 sampai 28 dengan jenis atribut numerik dan kategoris dan kombinasi keduanya, dan juga menggunakan data sintetik untuk mengetahui skalabilitasnya.

Hasilnya diperoleh bahwa algoritma *MIND* yang diimplementasikan ini terbukti memiliki tingkat akurasi yang lebih baik daripada *SPRINT* yaitu mampu menghasilkan akurasi 100% dimana tingkat akurasi tertinggi *SPRINT* hanya sampai 99,08%. Pohon keputusan yang dihasilkannya cukup baik dan mudah dimengerti. Namun waktu prosesnya lebih lambat daripada *SPRINT* hingga rata-rata 19,27 kali pada data dengan jumlah atribut banyak (mulai 20 atribut keatas). Untuk data banyak (dengan jumlah *record* 100.000 keatas), pengimplementasian algoritma *MIND* yang dibuat ini ternyata tidak mampu menghasilkan klasifikasi data sehingga hasil pengimplementasian ini belum *scalable* sebagai suatu *tools* pengklasifikasi data.

Kata kunci: data mining, algoritma *MIND*, klasifikasi, *scalable classifier*