

ABSTRAK

Data mining merupakan proses semi-otomatis untuk mengeksplorasi dan menganalisis data yang berjumlah besar gunanya untuk mendapatkan pola yang berguna. Data mining ini merupakan proses gabungan antar bidang-bidang terutama adalah *machine learning*, analisis statistik dan basis data. Data mining berusaha untuk menemukan kaidah dan pola dari data.

Salah satu *task* yang penting dalam data mining adalah *classification* (klasifikasi). Klasifikasi ini dapat dideskripsikan sebagai berikut: terdiri dari data input yang disebut juga sebagai *training set* terdiri dari sejumlah *examples* (*record*) yang masing-masing memiliki sejumlah atribut atau disebut juga fitur. Adapun tujuan klasifikasi ini adalah untuk menganalisis data input dan mengembangkan sebuah model yang akurat untuk setiap kelas berdasarkan beberapa variabel prediktor.

Melalui tugas akhir ini penulis akan mencoba mengimplementasikan sebuah algoritma berdasarkan *decision tree classifier* (pohon keputusan pengklasifikasi) yang melakukan fungsi klasifikasi pada data mining yakni algoritma SLIQ (Supervised Learning In Quest).

SLIQ adalah algoritma klasifikasi berdasarkan pohon keputusan yang dapat menangani atribut numerik maupun kategoris. SLIQ menggunakan kombinasi teknik *pre-sorting* pada fase pembangunan pohon, prosedur pengurutan yang terintegrasi dengan strategi pembuatan pohon *breadth-first* dan algoritma *pruning-tree* (pemotong pohon) baru berdasarkan prinsip *Minimum Description Length* (MDL). Kombinasi beberapa teknik ini menjadikan SLIQ sebuah algoritma yang dapat menangani data dalam ukuran besar sehingga penulis tertarik untuk mengangkatnya dalam tugas akhir ini.

Kata kunci : data mining, klasifikasi, SLIQ .