

Abstrak

Perkembangan dunia industri dan informasi secara tidak langsung menyebabkan semakin banyaknya data yang telah dikumpulkan. Data tersebut sebenarnya masih bisa dimanfaatkan. Salah satu cara untuk mengekstrak pola yang penting atau menarik dari sejumlah data yang sangat besar untuk menghasilkan informasi yang lebih berguna adalah dengan menggunakan *data mining*.

Salah satu metoda pada *data mining* adalah *classification* yaitu proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk memperkirakan kelas dari suatu objek yang labelnya tidak diketahui

Pada Tugas Akhir ini membahas salah satu algoritma pada *decision tree* yaitu Classification & Regresion Tree (C&RT atau CART) dan kemudian menganalisa performansinya. Algoritma C&RT memiliki karakteristik khusus yaitu selalu melakukan *splitting* terhadap *parent node* menjadi dua buah *child node*. Pengimplementasiannya dibantu dengan *tools* Clementine dan akan dibangun juga interface yang terintegrasi dengan *tools* Clementine, sehingga user mudah untuk melakukan dan melihat hasil klasifikasi.

Dari pengujian penggunaan algoritma C&RT yang telah dilakukan, didapatkan bahwa tingkat keakurasian sangat ditentukan oleh pemilihan data yang digunakan pada saat *training* dimana nantinya akan sangat menentukan *rule* untuk *testing*.

Kata kunci: *Data mining, Classification, Clementine, decision tree, C&RT, CART.*