

Abstrak

Data merupakan salah satu sumber yang digunakan untuk memperoleh suatu informasi. Namun tidak semua data dapat dimanfaatkan dengan baik. Jika data tersebut memiliki struktur yang kompleks, tentu saja akan sulit dimengerti. Sebagai contoh struktur data yang kompleks yaitu *microarray* yang digunakan pada Tugas Akhir ini. Data tersebut memiliki struktur dimensi yang besar dan *multi-label* sehingga akan sulit dipahami. Oleh karena itu diperlukan suatu proses *feature selection* yang bertujuan memperkecil ukuran data namun tetap menghasilkan nilai *accuracy* yang baik dan tidak terlalu besar jika terdapat penurunan terhadap nilai *accuracy* pada klasifikasi data.

Pada Tugas Akhir ini dilakukan perbandingan metode *feature selection* yaitu *Max-Relevance* dengan menggunakan beberapa *attribute evaluator* antara lain *GainRatioAttributeEval*, *InfoGainAttributeEval*, *Mutual Information*, dan *SymmetricalUncertAttributeEval*. Metode - metode *feature selection* tersebut diterapkan secara *filtering*. Proses klasifikasi data menggunakan bantuan tools *weka*, teknik klasifikasi yang digunakan ialah *Naïve Bayes*, dimana teknik ini memperhitungkan jumlah kemunculan data. Dengan melakukan analisis perbandingan terhadap metode *feature selection* maka dapat diketahui metode *feature selection* mana yang lebih handal dalam menangani data yang memiliki dimensi yang besar khususnya data *microarray*. Pengukur evaluasi yang dibandingkan ialah hasil proses klasifikasi yaitu *accuracy*.

Kata kunci: *microarray*, *feature selection*, klasifikasi, *naïve bayes*, *accuracy*.