

Abstrak

Data merupakan bahan utama dalam mendapatkan Informasi. Pengolahan data yang tepat dapat berimbas pada informasi bahkan pengetahuan yang berguna. Salah satu cara untuk mendapatkan pengetahuan dari suatu data adalah dengan proses klasifikasi. Proses klasifikasi sangat dipengaruhi oleh kondisi data. Kondisi data yang bersih bebas *noise* akan menyebabkan data dapat terklasifikasi dengan baik. Data dengan dimensi tinggi dapat dipastikan akan memiliki waktu proses yang sangat besar. Salah satu *noise* yang sering dijumpai adalah adanya *irrelevant* dan *redundant feature*.

Salah satu cara yang biasa digunakan untuk menyelesaikan permasalahan adanya *irrelevant* dan *redundant feature* adalah dengan menggunakan *feature selection*. salah satu metode *feature selection* yang cukup terkenal adalah *Fisher score*, tetapi *fisher score* sendiri memiliki beberapa kelemahan diantaranya adalah tidak dapat menangani *redundant feature* dan tidak dapat mengani kemungkinan terdapatnya subset *feature* pada suatu data.

Pada Tugas Akhir ini, akan dibahas sebuah metode yang menangani kelemahan dari metode *Fisher score*, yaitu *Generalized Fisher Score* dan dilakukan pula analisis terhadap peforma metode *Generalized Fisher Score* sebagai *feature selection*. Selain itu pada tugas akhir ini juga mencoba memperbaiki kelemahan yang ada pada *Generalized Fisher Score* dengan memberikan modifikasi pada *Generalized Fisher Score*.

Kata Kunci : *Data Mining*, Klasifikasi, *feature selection*, *Generalized Fisher Score*.