

Abstrak

Perkembangan teknologi memacu timbulnya keberagaman data didalamnya. Sedangkan data adalah sumber informasi yang sangat penting. Untuk dapat mengolah data-data tersebut terdapat teknik yang sekarang diimplementasikan, yaitu *Knowledge Discovery in Database* (KDD). Di dalam KDD terdapat proses *data mining* yang bertujuan untuk menggali informasi dari data yang ada. Salah satunya dengan cara klasifikasi. Akan tetapi dengan keberagaman data tidak ada jaminan bahwa data itu siap diolah. Contohnya adalah dimensi data yang begitu besar, hal ini akan menyulitkan dalam proses klasifikasi. Maka dari itu dilakukanlah *preprocessing* terlebih dahulu.

Preprocessing adalah tahap dimana untuk menyiapkan data agar seefisien mungkin dan terhindar dari *noise*, *missing value*, *irrelevant feature*, *redundant feature* dll, sehingga diharapkan akan memberikan hasil yang lebih optimal dalam melakukan klasifikasi. Di dalam *preprocessing*, terdapat salah satu teknik yaitu *feature selection*. Teknik ini digunakan untuk mengurangi dimensi data atau *feature* yang dianggap kurang relevan terhadap pembentukan kelas.

Tugas Akhir ini membahas serta mengimplementasikan teknik *feature selection* dengan menggunakan metode *Rough Sets Theory* yang dipadukan dengan *MLRelevance Criterion* dan *PRelevance Criterion*. Hasil dari *feature selection* dengan menggunakan metode itu, mampu memprediksi *feature* yang paling relevan. Sehingga tingkat akurasi yang didapatkan mampu mengimbangi *precision*, *recall* dan *accuracy* sebelum dilakukan *feature selection*.

Kata kunci : *Data mining*, *preprocessing*, klasifikasi, *Rough Set*, *feature selection*, *variable selection*.