

Abstrak

Clustering adalah suatu teknik dalam data mining yang bertujuan untuk membagi dan mengelompokkan suatu data berdasarkan kemiripan dari karakteristik pada data tersebut. Telah banyak algoritma untuk menerapkan *clustering* ini pada data numerikal. Namun pada data kategorikal, yaitu data yang bersifat kategoris dimana suatu objek memiliki *value* atau isi pada tiap atribut dimana tiap atribut memiliki himpunan *value*, masih belum banyak. Salah satu algoritma yang dapat menangani *clustering* pada data kategorikal adalah Min Min Roughness (MMR). MMR memanfaatkan teori rough set dan menggunakan satu parameter input berupa jumlah *cluster* yang ingin dibentuk, dimana teori rough set ini akan mendapatkan nilai *roughness* dari tiap *value* yang ada pada tiap atribut terhadap atribut lain. Dari nilai *roughness* pada tiap *value* pada suatu atribut, akan didapatkan nilai mean *roughness* nya yang selanjutnya akan digunakan dalam pemilihan atribut dan *value* dari atribut tersebut sebagai dasar dalam pemecahan data. Dari cluster-cluster yang terbentuk akan dianalisis tiap cluster tersebut dengan menggunakan metode *purity* dan berdasarkan objek-objek yang terdapat dalam cluster tersebut.

Kata Kunci : *Clustering*, MMR, Teori Rough Set