

Abstrak

Preprocessing data adalah proses pembersihan data dari *noise*, *outlier*, *missing value*, *irrelevant feature*, *redundant feature* dll agar data yang digunakan bisa lebih efisien saat dilakukan proses data mining selanjutnya. Salah satu cara *preprocessing* adalah dengan mereduksi dimensi data. Metode reduksi data yang bisa dilakukan adalah dengan memilih *feature-feature* yang penting saja (*feature selection*). Pada tugas akhir ini diterapkan metode *feature selection* MMIMRSC (*Maintaining mutual information and minimizing redundancy-synergy coefficient*) yang bertujuan menghilangkan *feature* irelevan dan redundan dengan menghitung nilai informasi *feature* serta mengurangi redundansi *feature* dengan menghitung *redundancy-synergy coefficient*nya. Hasil performansi data hasil *preprocessing*, yaitu nilai *precision* dan *recall* nya mengalami peningkatan dan waktu pembentukan model klasifikasi menjadi lebih cepat.

Kata kunci : *feature selection*, *mutual information*, reduksi data, *preprocessing*