

Abstrak

Missing value merupakan salah satu penyebab suatu data dikatakan tidak valid atau masih merupakan data 'kotor'. *Missing value* dengan jumlah banyak akan menghilangkan sejumlah informasi yang seharusnya dibutuhkan untuk diolah. Salah satu metode penanganan yang dapat dilakukan pada *missing value* adalah dengan mengisi suatu nilai kepada data-data yang *missing*. Nilai untuk mengisi *missing value* diperoleh dengan prediksi dari informasi yang masih tersedia pada data. Metode ini disebut dengan metode imputasi. Salah satu metode imputasi yaitu *K-Means Imputation* (KMI). Metode ini adalah metode yang didasarkan pada algoritma *K-Means* yang biasa dipakai dalam proses *clustering*. KMI akan melakukan pengisian *missing value* dengan menghitung rata-rata nilai atribut dari data-data yang berada pada suatu cluster dimana data missing juga berada dan data-data tersebut haruslah memiliki kelas yang sama dengan yang dimiliki data missing. Pada tugas akhir ini, akan dicoba untuk melakukan imputasi missing value dengan menggunakan *K-Means Imputation* (KMI). Parameter yang akan diuji adalah *root mean square error*, *precision*, *recall* dan *f-measure*. Berdasarkan hasil pengujian, KMI terbukti dapat memprediksi nilai missing value mendekati nilai data yang sebenarnya pada data yang memiliki tingkat keragaman variasi nilai atribut kecil dan dapat meningkatkan hasil klasifikasi.

Kata Kunci: metode imputasi, *K-Means Imputation* (KMI), *missing value*.