

Abstrak

Salah satu permasalahan yang sering dimiliki oleh data mentah adalah adanya *missing value*. Hal ini tentunya akan mempengaruhi hasil analisis statistik pada data sehingga dapat mengurangi tingkat keakuratan data. Oleh karena itu dibutuhkan tahap preprocessing data untuk dapat menangani masalah *missing value* ini. Salah satu cara penanganannya adalah dengan mengisi nilai prediksi yang diperoleh dengan menerapkan suatu algoritma tertentu, yang disebut dengan metode imputasi. Metode imputasi *missing value* yang digunakan dalam Tugas Akhir ini adalah metode *Predictive Mean Matching* (PMM).

PMM merupakan metode imputasi yang menerapkan algoritma yang berbasis model, dimana *missing value* diisi oleh nilai estimasi yang diperoleh dari model menggunakan *Least Square Regression* yang terbentuk dari baris yang tidak memiliki *missing value* pada data kemudian mengimputasi *missing value* dari nilai observasi yang terdekat dari model. Pengujian performansi hasil imputasi pada sistem dievaluasi melalui parameter *Mean Squared Error* (MSE). Pada Tugas Akhir ini juga dilakukan analisis pengaruh imputasi data terhadap proses klasifikasi dengan parameter pengujian *Precision*, *Recall*, dan *F-Measure*. Berdasarkan hasil pengujian yang dilakukan, terbukti bahwa PMM merupakan algoritma imputasi yang dapat mengestimasi nilai *missing value* yang mendekati nilai sebenarnya.

Kata kunci: imputasi, *missing value*, *Predictive Mean Matching* (PMM)