

Abstract

Missing value is one of most common problem in machine learning which is the missing information from data for specific reasons. Missing value in small quantities does not significantly affect the data, but in a large quantities it can reduce the accuracy of data. Therefore it is necessary to fill in missing value approach to the real value, was called the imputation method. One of the imputation method that being adopted in this Final Project is Collateral Missing Value Estimation method (CMVE).

This method applies the principle of multiple imputation to estimate the possible values. The estimated value is calculated using Least Square Regression and Non-Negative Least Square. Then the value were given by combining that two method with weight equal to one as final estimated value. Normalized Root Mean Square Error (NRMSE) was used as the evaluation performance of this imputation system. *Precision*, *recall*, dan *f-measure* were used to evaluated the performance in classification. Based on the test CMVE method is able to generate imputation value approach to the real value.

Keywords: imputation, missing value, *Collateral Missing Value Estimation* (CMVE)