ABSTRACT

Data is one most important key in researching. The quality in one research will be highly affected by the quality of the data that researcher collect. One most deadly problem is missing data in the data set. One of many way to find missing data is k-Nearest Neighbour(KNN). But this method have a weakness. That is, k value that researcher don't know what number to set it. This project goal is to find a way to search optimal k value to input on KNN. Memetic algorithm is the metod that used in this journal project to find that. This project try to find optimal k value on KNN with memetic algorithm, that can minimize MSE as mush as possible. Performance testing is done by comparing the value of MSE from the imputation of missing data when there were missing data, 5%, 10%, 15%, 20% and 25% with KNN method and KNN with memethic algorithm optimation. Results from the study showed that the k obtained from the memetic algorithm optimization does not necessarily indicate the optimum k with the best accuracy measurement results are tested, but very close to the best results of the measurement accuracy is tested.

Keywords : missing data, KNN, memetic algorithm, optimation