Abstract

In database where missing data becomes more important, data prediction is one of the widely used approach that seeks to predict missing data in a data set. Many researches implement models to predict missing data. But the results are of low accuracies. The Multiple Imputation Model developed in this study predicts more accurately for missing values in different location. Previous researches usually focused only either on the average value or similarity value consequently low accuracy occurs. To get more accurate data prediction in a data set, additional factors such as data types, personal preference and relationship of variables need to be considered.

In this study where EachMovie data set was used to evaluate the validity and accuracy of the Multiple Imputation Model. Evaluation using Mean Absolute Error (MAE) revealed that data grouping of the EachMovie data set into movie types such as Animation, Romance, Action, Comedy, Art_Forein, Classic, Drama, Family, Horror, and Thriller did not give accurate prediction. Data Clustering gave more accurate results.