Privacy Preserving in Data Mining by Normalization

Authors:
Syed Md. Tarique Ahmad
Shameemul Haque
Prince Shoeb Khan

10.5120/16797-6509

Abstract

Extracting previously unknown patterns from massive volume of data is the main objective of any data mining algorithm. In current days there is a tremendous expansion in data collection due to the development in the field of information technology. The patterns revealed by data mining algorithm can be used in various domains like Image Analysis, Marketing and weather forecasting. As a side effect of the mining algorithm some sensitive information is also revealed. There is a need to preserve the privacy of individuals which can be achieved by using privacy preserving data mining. In this paper we use min- max normalization approach for preserving privacy during the mining process. We clean the original data using min- max normalization approach before publishing. For experimental purpose we have used k- means algorithm and from our results it is obvious that our approach preserves both privacy and accuracy.

References


Index Terms

Computer Science  Data Mining

Keywords

Clustering  K- Means  Accuracy  Privacy  Min-Max Normalization  Normalization.