

Abstract

The well-known privacy-preserved data mining modified existing data mining techniques to randomized data. In this project, data mining is use as a technique for masking data, therefore, termed data mining based privacy protection. This approach incorporates partially the requirement of a targeted data mining task into the process of masking data so that essential structure is preserved in the masked data. Bottom-up generalization concept is use as a way to hide detailed information, rather than discover trends and patterns. Ince data is masked, standard data mining techniques can be applied without modification. This project demonstrated another positive use of data mining technology, not only can discover useful patterns, but also mask private information. Testing is done to see whether the results generalize the data still fit for use for further data mining task is classification. The accuracy of the data changed after the data being protected, but not so large that the data can still be used for classification.

Keyword: privacy-preserving data mining, bottom-up generalization, classification