Abstract

Data Mining has many applications in the real world. One of the most important and widely found problems is that of classification. Recently, distance preserving data perturbation has gained attention because it mitigates the privacy/accuracy trade-off by guaranteeing perfect accuracy. Many important data mining algorithms can be efficiently applied to the transformed data and produce exactly the same results as if applied to the original data. e.g., distance-based clustering and k-nearest neighbor classification. In this research paper we analysis Euclidean distance-preserving data perturbation for k-nearest neighbor classification as a tool for privacy-preserving data mining.

References


Index Terms

Computer Science
Artificial Intelligence

Keywords

K-nearest neighbor classification