Abstract

Privacy is one of the most concerned issues in data publishing. Personal data like electronic health records and financial transaction records are usually deemed extremely sensitive although these data can offer significant human benefits if they are analyzed and mined by organizations such as disease research Centre. The emerging research field in data mining, Privacy Preserving Data Publishing (PPDP) [11], targets these challenges. The basic idea of PPDM is to modify the data in such a way so as to perform data mining algorithms effectively without compromising the security of sensitive information contained in the data. It aims at developing techniques that enable publishing data while minimizing data distortion for maintaining utility and ensuring that privacy is preserved.

References


Index Terms

Computer Science Software Engineering

Keywords

privacy, privacy preserving data mining (ppdm), k-anonymity, suppression, generalization, mapreduce.