

STATISTICAL ANALYSIS AND DATA MINING

Original Article

Interactive knowledge discovery from hidden data through sampling of frequent patterns*

Mansurul Bhuiyan  Mohammad Al Hasan

First published: 13 July 2016

<https://doi.org/10.1002/sam.11322>

Cited by: 2

* This manuscript is an extended version of a paper that was published in CIKM (International Conference on Information and Knowledge Management) 2012 conference. The pdf copy of the conference paper is available from <http://www.cs.iupui.edu/alhasan/papers/fp006 - bhuiyan.pdf>.



About



Access

ols



ire

Abstract

In real life, many important datasets are not publicly accessible due to various reasons, including privacy protection and maintenance of business competitiveness. However, Knowledge discovery and pattern mining from these datasets can bring enormous benefit both to the data owner and the external entities. In this paper, we propose a novel solution for this task, which is based on Markov chain Monte Carlo (MCMC) sampling of frequent patterns. Instead of returning all the frequent patterns, the proposed paradigm sends back a small set of randomly selected patterns so that the confidentiality of the dataset can be maintained. Our solution also allows interactive sampling, so that the sampled patterns can fulfill the user's requirement effectively. We show experimental results from several real - life datasets to validate the capability and usefulness of our solution. In particular, we show examples that by using our proposed solution, an eCommerce marketplace can allow pattern mining on user session data without disclosing the data to the public; such a mining paradigm can help the sellers in the marketplace, which eventually can boost the market's own revenue. © 2016 Wiley Periodicals, Inc. Statistical Analysis and Data Mining: The ASA Data Science Journal, 2016

Citing Literature



About Wiley Online Library

Help & Support

Opportunities

Connect with Wiley



Copyright © 1999-2018 John Wiley & Sons, Inc. All rights reserved

WILEY