Abstract

Data mining is an important technology for extracting useful information from large collections of data in the database. Data mining techniques like classification rule mining and automated data collections have given the way to making automated decisions for loan granting or denial, personal selection etc. If the training data sets are biased with discriminatory or sensitive attributes like gender, age, religion, color etc., discriminatory decisions may ensue. That cause potential privacy invasion and potential discrimination. Later one consists of unfairly treating people on the basis of their belonging to a specific group. The anti-discrimination techniques named discrimination discovery and prevention have been introduced in data mining to solve these problems. Discrimination is divided into two, Direct and Indirect and it tackles discrimination prevention in data mining and propose new techniques applicable for direct, indirect and both at the same time. It also describes how to clean training data sets and outsourced data sets in such a way that direct and/or indirect discriminatory decision rules are converted to legitimate (nondiscriminatory) classification rules and a number of papers mention measures of utility too. This survey paper is aimed at understand the existing discrimination prevention techniques and the utility measures discussed so far.
A Survey on Direct and Indirect Discrimination Prevention in Data Mining

References

- Dino Pedreschi, Salvatore Ruggieri and Franco Turini &quot;Discrimination-aware Data Mining&quot; Dipartimento di Informatica, Università di Pisa, 2008 ACM, Las Vegas, Nevada, USA
- Faisal Kamiran, Asim Karim, and Xiangliang Zhang &quot;Decision Theory for Discrimination-aware Classification&quot; King Abdullah University of Science and Technology (KAUST), The Kingdom of Saudi Arabia, IEEE 12th International Conference on Data Mining 2012.
- Faisal Kamiran and Toon Calders &quot;Classification with No Discrimination by Preferential Sampling&quot; Eindhoven University of Technology, Netherlands, 19th Machine Learning conference of Belgium and The Netherlands. 2010
- Toon Calders and Sicco Verwer &quot;Three naive Bayes approaches for discrimination-free classification&quot;, This article is published with open access at Springerlink.com, 2010
- Sara Hajian, Josep Domingo-Ferrer and Antoni Mart`e &quot;Discrimination Prevention in Data Mining for Intrusion and Crime Detection&quot; Universitat Rovira i VirgiliDept. of Computer Engineering and Maths, UNESCO Chair in Data Privacy, Tarragona, Catalonia, 2011 IEEE
- S. Hajian and J. Domingo-Ferrer. &quot;A methodology for direct and indirect discrimination prevention in data mining&quot; 2012 IEEE
- P. N. Tan, M. Steinbach, and V. Kumar, &quot;Introduction to Data Mining&quot; Addison-Wesley, 2006.
- Sara Hajian, Josep Domingo-Ferrer, and Antoni Mart`e &quot;Rule Protection for Indirect Discrimination Prevention in Data Mining&quot; Springer 2011

Index Terms

Computer Science Information Sciences

Keywords

Data mining Discrimination Prevention Direct and Indirect Discrimination