Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 5, May 2014, pg.534 – 539

RESEARCH ARTICLE

Inference Mining using Direct and Indirect Discrimination Prevention in Data Mining

S. Shanthi[1], A. Gomathi[2]

[1]Faculty of Engineering, Department of Computer Science and Engineering
Anna University, Chennai, India
[2]M.E., Scholar, Department of Computer Science and Engineering
Anna University, Chennai, India

[1] shan.sece@gmail.com; [2] gomathianandhan2007@gmail.com

***Abstract* — Data Mining is an essential and flourishing technology to extract the relevant and useful information hidden in the large collections of data. Privacy preservation in data mining is an important issue when considering the legal and ethical aspects of data mining. Discrimination is one of the facts that pave the way for negative perceptions in the data mining. Direct and Indirect discrimination are the two types of discrimination which involves in sensitive and non sensitive attributes. In this paper, our effort is to prevent the discrimination from both types and proposing the inference mining technique which allows the decision making process whether to accept or deny the request.**

***Keywords*— DD-Direct Discrimination; IDD-Indirect Discrimination; PPDM-Privacy Preserving Data Mining; IM-Inference Mining; PD-Potential Discrimination; PND-Potential Non Discrimination**

I. INTRODUCTION

Data Mining is an important technology which is used to extract the relevant information from the huge collections of data. While considering the ethical and legal aspects, privacy which is an important factor in the data mining. Discrimination emerges when the privacy preservation lacks the security. Discrimination is unfair or unequal treatment of people based on membership to a category or a minority without regard to individual merit. Potential discrimination takes place on the basis of their specific group. There are however more obvious, every individual do not want to be discriminated because of their gender, religion, nationality, age and so on., especially when those attributes are used for making decisions about them like giving them a job, loan, insurance etc., For example, the European union implements the principle of equal treatment between men and women in the access to and supply of goods and services in matters of employment and occupation. Many antidiscrimination techniques regarding discrimination discovery and discrimination prevention have been introduced in data mining. Discrimination is of two types (i.e.) Direct and Indirect. Sensitive attributes are taken into account in the case of Direct which is otherwise known as Potential Discrimination (PD). Whereas non sensitive attributes is strongly correlated with the biased sensitive ones. Technology that provides the services to allow for automatic and routine collection of huge amount of data at first sight, automatic decisions may appear a sense of fairness. Classification rules do not direct by personal preferences. Eventually at a closer look, one could realise that the rules are in fact learned by the system from the data set.

Discrimination may be in direct and indirect way. DD consists of rules and procedures that directly mention minority or disadvantaged groups based upon the sensitive discriminatory attributes. IDD occurs when an organization makes a decision or puts in place a particular policy or practice, which, on the face of it appears to treat everyone equally, but which actually, in practise leads to people from a protected group being treated less favourable than other people. Actually IDD does not explicitly mention the rules and procedures but moreover it is strongly correlated with the biased sensitive ones.

1.1 Understanding discrimination

Throughout history, many people have experienced discrimination. Discrimination occurs as a result of prejudice, or bias. Prejudice is an opinion about a certain group of people based on a single characteristic, such as ethnicity, religion, gender, or age. This kind of prejudice that causes discrimination has at its root an unfavourable opinion which is based on a lack of knowledge and understanding.

Fig. 1 Understanding Discrimination

Discrimination covers any scenario where an individual or company treats a person or group of people unfairly due to a particular characteristic. The most commonly referred to types of discrimination include prejudicial behaviour that is based on skin colour, gender, disability, religion or sexual orientation. However, discrimination can be based on any attribute that is viewed as marking the victim out as being different. For example, discrimination might be on the basis of age, weight, height, nationality or any number of other attributes. In order to tackle this hurtful and damaging behaviour, it's important to identify the causes of discrimination. Earnings differentials or occupational differentiation is not in and of itself evidence of employment discrimination. Discrimination can be intended and involves disparate treatment of a group or unintended, yet create disparate impact for a group. In neoclassical economics theory, labour market discrimination is defined as the different treatment of two equally qualified individuals on account of their sensitive attributes. Racial discrimination differentiates individuals on the basis of real and perceived racial differences and has been official government policy in several countries. Regional or geographic discrimination based on the region in which a person lives or was born. It differs from national discrimination in that it may not be based on national borders or the country the victim lives in, but is instead based on prejudices against a specific region of one or more countries. Diversity of language is protected and respected by most nations who value cultural diversity. However, people are sometimes subjected to different treatment because their preferred language is associated with a particular group, class or category. Commonly, the preferred language is just another attribute of separate ethnic groups. Discrimination exists if there is prejudicial treatment against a person or a group of people who speak a particular language or dialect.

Discrimination discovery is about finding about discriminatory decisions hidden in a dataset of historical decision records. The basic problem in the analysis of discrimination given a dataset of historical decision records is to quantify the degree of discrimination suffered by a given group in a given context with respect to the classification decision.

II. PROPOSED SYSTEM

Classical system tackles the discrimination prevention in data mining and proposed new techniques applicable for both direct and Indirect discrimination prevention individually or both at the same time. Also, it mainly focuses to prevent discrimination technique of direct and indirect and lacks in the learning mechanism. In our proposed system, the discrimination data sets are identified and the prevention has been taken by means of the binary and categorical attributes. Inference Mining provides learning mechanism of every data sets. The system has five major modules and each module has its own features.

2.1 Basic Definitions

- A data set is a collection of data objects (records) and their attributes.
- An item is an attribute along with its value, e.g., Race=black.
- An item set i.e., X is a collection of one or more items, e.g., {foreign worker = yes, City= AVG}.
- A classification rule is an expression $X \rightarrow C$, where C is a class item (a yes/no decision), and X is an itemset containing no class item, e.g., {Foreign worker =Yes; City = NYC} → Hire = no. X is called the premise of the rule.
- The support of an item set, supp(X), is the fraction of records that contain the item set X. We say that a rule $X \rightarrow C$ is completely supported by a record if both and C appear in the record.
- The confidence of a classification rule, $conf(X \rightarrow C)$ measures how often the class item C appears in records that contain X. Hence, if $supp(X) > 0$ then
- $conf(X \rightarrow C) = \frac{supp(X,C)}{supp(X)}$

2.2 Preprocessing

Preprocessing is needed to make the data more suitable for data mining. Original data set are separated into direct as well as indirect discrimination based on the sensitive and non sensitive attributes as shown in the fig.2. Redundant and noisy data are eliminated and the cleaned data sets are ready to transform. Potential discrimination or the sensitive attributes are analysed and taken it as the direct one and the non potential discrimination attributes are identified and consider it as the indirect one. Sensitive attributes are race, religion, gender, nationality, disability, Marital status and age. Non sensitive attributes are thickly related to the sensitive one, along it would generate the direct discrimination.

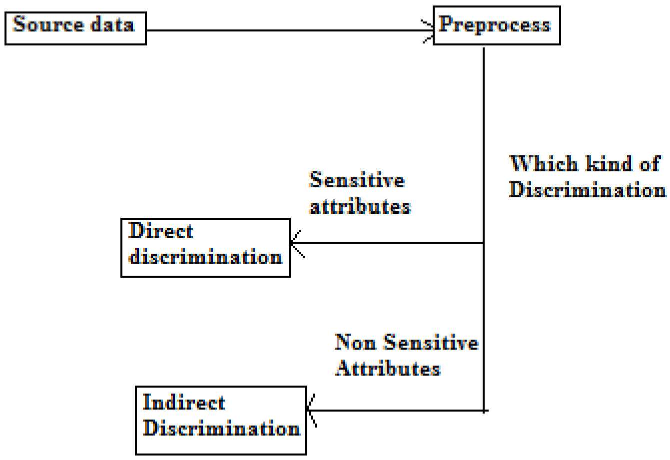

Fig 2. Preprocessing

2.3 Direct Discrimination

2.3.1 α-discrimination

α-discrimination is a fixed threshold stating an acceptable level of discrimination according to laws and regulations. Petreschi et al. Proposed a measure called extended lift (elift) which is based upon PD rule. Let A,B → C be a classification rule such that conf(B→ C) > 0. The extended lift of the rule is

$$\text{elift}(A,B\to C) = \frac{\text{conf}(A,B\to C)}{\text{conf}(B\to C)}$$

The idea here is to evaluate the discrimination of a rule as the gain of confidence due to the presence of the discriminatory items (i.e., A) in the premise of the rule. Whether the rule is to be considered discriminatory can be assessed by thresholding elift as follows.

Let $\propto \in R$ be a fixed threshold and let A be a discriminatory item set. A PD classification rule A,B → C is $\propto$ protective w.r.t. elift if elift(c) < $\propto$. Otherwise, c is $\propto$-discriminatory.

The purpose of direct discrimination discovery is to identify $\propto$ discriminatory rules. In fact, $\propto$-discriminatory rules indicate biased rules that are directly inferred from discriminatory items (e.g., Foreign worker = Yes). We call these rules direct $\propto$ discriminatory rules.

Direct discrimination is based on the concept of classification without discrimination which is massaging the data attributes to remove the discrimination from it with the least possible changes. The changes are made by changing the class labels of the selected attributes in the training data set in order to obtain a discrimination free set.

2.4 Indirect Discrimination

2.4.1 Redlining rules

Redlining is the practice of denying, or charging more for, services such as banking, insurance, access to health care, or even supermarkets, or denying jobs to residents in particular, often racially determined areas. Redlining by financial services (denying to grant mortgages or insurances in urban areas considering it as deteriorating) is an archetypal example of indirect discrimination. Indirect discrimination is based upon the categorical attributes in which the data sets are modified as a class label which becomes the discrimination free set.

A PND classification rule r : D;B → C is a redlining rule if it could yield an α-discriminatory rule r' : A,B → C in combination with currently available background knowledge rules of the form rb1 : A,B → D and rb2 : D,B → A, where A is a discriminatory item set. For example, {Zip = 10451; City = NYC} → Hire = No.

A PND classification rule r : D,B → C is a non redlining or legitimate rule if it cannot yield any α- discriminatory rule r' : A,B → C in combination with currently available background knowledge rules of the form rb1 : A;B → D and rb2 : D;B → A, where A is a discriminatory item set. For example, {Experience = Low; City = NYC} → Hire = No. We call α-discriminatory rules that ensue from redlining rules indirect α-discriminatory rules.

2.5 Inference Mining

2.5.1 Algorithm

1: Inputs: DB, FR, α, DIs
2: Output: DB´ (transformed data set)
3: for each DIs,
4: compute α
5: elift = conf(A,B→C)
6: Find α protective
7: DataTransformation
8: while Direct r : DI's→c or c'
9: end while
10: Redlining rules
11: while Indirect r : DI's→categorical
12: end while
13: inference r: supp (DD, IDD)
14: compare threshold value
15: Transformed data sets

Based upon the measures of interestingness the threshold value is assigned and upon that the proposal is whether accepted or deny. As we know the support and confidence are the two measures of interestingness.

III. CONCLUSION

Along with privacy, discrimination is an important issue when considering the legal and ethical aspects of data mining. The objective of this proposed system is to develop prevention over the discrimination either PD or PND. To attain this objective initially, the data attributes are preprocessed and separated into direct and indirect. The DD is measured on elift and class labels are changed in to binary attributes. The IDD are using the background knowledge i.e. the redlining rules and are slightly modified into the categorical attributes. Finally the inference mining is using the technique similar to Association rule mining and with the threshold value the proposal may accept or not accept.

REFERENCES

1. Calders T. and Verwer S., (2010) "Three Naive Bayes Approaches for Discrimination-Free Classification," Data Mining and Knowledge Discovery, vol. 21, no. 2, pp. 277-292.
2. Hajian S. Domingo-Ferrer J, and Martı´nez-Balleste´A.,(2011) "Discrimination Prevention in Data Mining for Intrusion and Crime Detection," Proc. IEEE Symp. Computational Intelligence in Cyber Security (CICS '11), pp. 47-54.
3. Hajian S., Domingo-Ferrer J, and Martı´nez-Balleste´A., (2011) "Rule Protection for Indirect Discrimination Prevention in Data Mining," Proc. Eighth Int'l Conf. Modeling Decisions for Artificial Intelligence (MDAI '11), pp. 211-222.
4. Kamiran F. and Calders T, (2010) "Classification with no Discrimination by Preferential Sampling," Proc. 19th Machine Learning Conf. Belgium and The Netherlands.
5. Kamiran F., Calders T, and Pechenizkiy M, (2010) "Discrimination Aware Decision Tree Learning," Proc. IEEE Int'l Conf. Data Mining (ICDM '10), pp. 869-874.
6. Pedreschi D., Ruggieri S., and Turini F., (2008) "Discrimination-Aware Data Mining," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (KDD '08), pp. 560-568.

7. Pedreschi D., Ruggieri S., and Turini F., (2009) "Measuring Discrimination in Socially-Sensitive Decision Records," Proc. Ninth SIAM Data Mining Conf. (SDM '09), pp. 581-59.

8. Pedreschi D., Ruggieri S., and Turini F., (2009) "Integrating Induction and Deduction for Finding Evidence of Discrimination," Proc. 12th ACM Int'l Conf. Artificial Intelligence and Law (ICAIL '09), pp. 157-166.

9. Ruggieri S., Pedreschi D., and Turini F., (2010) "Data Mining for Discrimination Discovery," ACM Trans. Knowledge Discovery from Data, vol. 4, no. 2, article 9.

10. Ruggieri S., Pedreschi D., and Turini F., (2010) "DCUBE: Discrimination Discovery in Databases," Proc. ACM Int'l Conf. Management of Data (SIGMOD '10), pp. 1127-1130.

