International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 11, November - 2013

Review on Discrimination Prevention Using both Direct and Indirect Method in Data Mining

Sampada U. Nagmote
P.R.Patil College of Engg & Technology,
Amravati

Prof. P. P. Deshmukh
P.R.Patil College of Engg & Technology,
Amravati

Abstract

Today, Data mining is an increasingly important technology. It is a process of extracting useful knowledge which is in large collections of data. There are some negative understanding about data mining, among which potential privacy invasion and potential discrimination. Discrimination means is the unequal or inferior treatment of persons on some of certain characteristics or unfairly treating people on the basis of their specific belonging group. If the data sets are divided on the basis of sensitive attributes like gender, race, religion, etc., discriminatory decisions may ensue. For this reason, antidiscrimination techniques for discrimination prevention have been introduced for data mining. Discrimination can be either direct or indirect. Direct discrimination occurs when decisions are made based on some sensitive attributes. It consists of rules or procedures that explicitly mention minority or disadvantaged groups based on sensitive discriminatory attributes related to group membership. Indirect discrimination occurs when decisions are made based on nonsensitive attributes which are strongly related with biased sensitive ones. It consists of rules or procedures that, which is not explicitly mentioning discriminatory attributes, intentionally or unintentionally, could generate decisions about discrimination. In this paper, we discuss basic definition of direct and indirect discrimination, discrimination prevention in data mining, rule protection etc.

1. Introduction

Discrimination is an act of unfairly treating people on the basis of their belonging to some aspecific group. For instance, individuals may be discriminated because of their race, gender, etc. [5] or it is the treatment to an individual based on their membership in particular category or group. There are various laws which are used to prevent discrimination on basic of various attributes such as race, religion, nationality, disability and age.

There are two types of discrimination i.e. Direct Discrimination and Indirect Discrimination. Direct Discrimination is direct discrimination which consists of rules or procedure that mention minority or disadvantaged group based on sensitive attributes to they are related to membership of group. Indirect Discrimination is discrimination which consists of rules and procedures that are not mentioning attributes and hence it generates discriminatory decision intentionally or unintentionally. [1]

The points which will discuss in this paper are:-

- Process for Discrimination Discovery
- Basic Definitions
- Direct Discrimination Prevention Method
- Indirect Discrimination Prevention Method

2. Process for Discrimination Discovery

Process of discrimination discovery is about for finding out discriminatory decisions which are in a dataset. The most basic problem in the discrimination analysis, given a dataset , is to quantify the degree of discrimination suffered by a given group in a given context with respect to the classification decision.

Figure 1 shows the process of discrimination discovery, based on approaches and measures described in this section. [5]

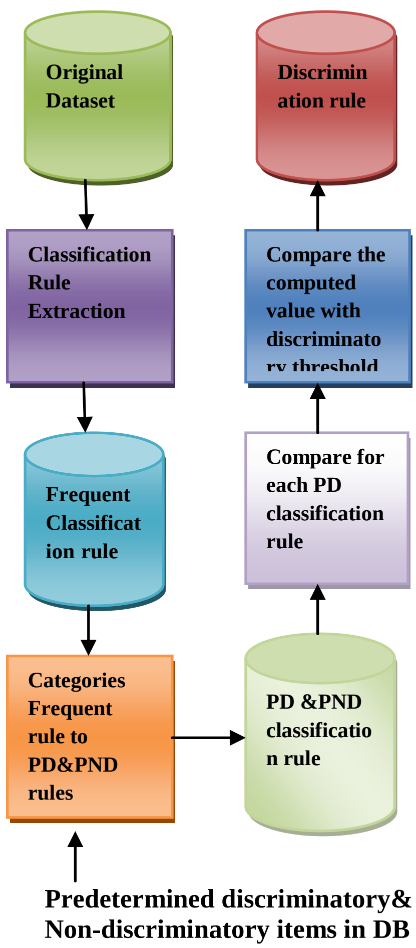

Fig1. Process of Discrimination Discovery

3. Basic Definition

- An item is an attribute with its value, e.g.{Gender =Female}.
- In association/ classification rule mining ,it attempts, to predict the occurrence of an item based on the occurrences of other items in the transaction from the given a set of transactions (records).
- An item set is a set of one or more items, e.g. {Gender=Male, Zip=54341}.
- A classification rule is an expression X →C, whereas X is an set of item which containing no class items, and C is a class item, e.g.{Gender=Female, Zip=54341} → Intruder=YES.X is called the premise (or the body) of the rule.
- The support of an itemset, supp(X), is the fraction of records that contain the item set X. We say that a rule X → C is completely supported by a record if both X and C appears in the record.
- The confidence of a classification rule, Conf(X →C), measures how often the class item C appears in records that contain X. Hence, if supp(X) > 0

$$\textbf{Conf}\ (\mathbf{X} \rightarrow \mathbf{C}) = \frac{\textbf{supp(X, C)}}{\textbf{supp (X)}} \qquad (1)$$

Support and confidence range over [0, 1].

- In addition, the notation also extends to negated item sets, i.e. ¬X.
- A frequent classification rule is a classification rule with a support or confidence greater than a specified lower bound. Let DB be a database of original data records and FRs be the database of frequent classification rules.[5]

4. Direct Discrimination Prevention

4.1 Definition of Direct Discrimination

Definition 1. Let A; B → C be a classification rule such that conf(B→C) > 0. The extended lift of the rule is

$$\text{elift}\ (A;B\rightarrow C) = \frac{\text{conf}\ (A;B\rightarrow C)}{\text{conf}(B\rightarrow C)} \qquad (2)$$

Here the idea is that to evaluate the discrimination of a rule.

Definition 2. Let α € R be a fixed threshold[1] and let A be a discriminatory item set. A PD classification rule c =A, B →C, C is α-protective w.r.t. elift if elift(c) < α. Otherwise, c is α-discriminatory. The aim of direct discrimination discovery is to identify α -discriminatory rules. [12][13]

4.2 Direct Rule Protection

In order to convert each α-discriminatory rule into an α-protective rule, which based on the direct discriminatory measure (i.e., Definition 2), we should have the following inequality for each α-discriminatory rule r': A;B→C in MR, where A is a discriminatory item set:

$$\text{elift}\ (r') < \alpha \qquad (3)$$

This inequality can be rewritten as

$$\frac{\text{Con}f(r'\text{: A, B}\rightarrow\text{C})}{\text{Con}f\ (\text{B}\rightarrow\text{C})} < \alpha \quad (4)$$

Let us rewrite Inequality (4) in the following way:

$$\text{Con}f(r'\text{: A, B}\rightarrow\text{C}) < \alpha.\ \text{Con}f\ (\text{B}\rightarrow\text{C}) \quad (5)$$

So, it is clear that Inequality (3) can be satisfied by decreasing the confidence of the α-discriminatory rule r': A; B →C to a value less than the right-hand side of Inequality, without affecting the confidence of its base rule B→C. Then possible solution for decreasing is

$$\text{Con}f(r'\text{: A, B}\rightarrow\text{C}) = \frac{\text{supp (A, B, C)}}{\text{supp (A,B)}} \qquad (6)$$

This is to perturb the discriminatory item set from ¬A to A in the subset DBc of all records of the original data set which completely support the rule r': A;B→ C and have some impact on other rules; for doing this increases the denominator of Expression (6) while keeping the numerator and Con$f$ (B→C) as it is. There is also another way to provide direct rule protection.
Let us rewrite above Inequality in the following different way:

$$\text{con}f\ (\text{B}\rightarrow\text{C}) >= \frac{\text{con}f\ (r'\text{: A, B}\rightarrow\text{C})}{\alpha} \qquad (7)$$

It is clear that Inequality (3) can be satisfied by increasing the confidence of the base rule (B→C) of the α-discriminatory rule r': A; B→C to a value higher than the right-hand side of Inequality (7), without affecting the value of con$f$(r': A; B→C). A possible solution for increasing Expression

$$\text{con}f\ (\text{B}\rightarrow\text{C}) \quad \frac{\text{supp (B, C)}}{\text{supp (B)}} \qquad (8)$$

is to perturb the class item from ¬C to C in the subset DB_c all records of the original data set which completely support the rule ¬A;B→¬C and have minimum impact on other rules; doing so increases the numerator of Expression (8) while keeping the denominator and con$f$(r': A,B →C)
There are some other methods that could be applied for direct rule protection.

Similar data transformation methods could be applied to obtain direct rule protection with respect to other measures (i.e., slift and olift). [1] Algorithms based on this rule protection are given in paper [1].

5. Indirect Discrimination Prevention

5.1 Definition of Indirect Discrimination

Definition1. A PND classification rule r:
X (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 rb: A, B→ D and rb2: D, B →A, whereas A is a discriminatory itemset.

Definition2. A PND classification rule r : X(D,B)→C is a non-redlining 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 itemset.
In correlation between the discriminatory itemset A and the non-discriminatory itemset D with context B indicated by the background rules rb1and rb2 holds with confidences at least β1 and β2, respectively; however, it is not a completely certain correlation.
Let RR be the database of redlining rules extracted from database DB. [6]

5.2 Indirect Rule Protection

In order to turn out redlining rule into an nonredlining rule, based on the indirect discrimination measure, we should enforce the following inequality for each redlining rule
r: D; B →C in RR:

$$\text{elb}\ (\gamma, \delta) < \alpha. \qquad (9)$$

The above inequality can be rewritten as

$$\frac{\frac{\text{Con}f\ (\text{rь1})}{\text{Con}f\ (\text{rь2})}\ (\text{con}f\ (\text{rь2}) + \text{con}f(\text{r: D, B}\rightarrow\text{C})\text{-}1)}{\text{con}f\ (\text{B}\rightarrow\text{C})} < \alpha\ (10)$$

Discriminatory item set (i.e., A) is not removed from the original database DB and the rules rb1: A; B →D and rb2: D; B →A are obtained from DB, so that their Confidences might change and which is result of data transformation for indirect discrimination prevention.
Let us rewrite the above inequality in the following way:

$$\alpha.\ \text{Con}f\ (\text{B}\rightarrow\text{C}).\ \text{Con}f\ (\text{rb2})$$

$$\text{Con}f\ (\text{rb1: A, B}\rightarrow\text{D}) < \frac{\qquad}{\text{Con}f\ (\text{rb2}) + \text{Con}f(\text{r: D, B}\rightarrow\text{C})\text{-1}} \tag{11}$$

Clearly, in this case inequality (9) can be satisfied by decreasing the Confidence of rule rb1: A; B→ D to values less than the right-hand side of Inequality (11) without affecting either the Confidence of the redlining rule or the confidence of the B→ C and rb2 rules. Since the values of both inequality sides are dependent, a transformation is required that decreases the left-hand side of the inequality without any impact on the right-hand side. A possible solution for decreasing

$$\text{Con}f\ (\text{A, B}\rightarrow\text{D}) = \frac{\text{supp (A, B,D)}}{\text{supp (A, B)}} \tag{12}$$

in Inequality (11) to the target value is to perturb the discriminatory item set from ¬A to A in the subset DBc of all records of the original data set which completely support the rule ¬A; B; ¬D→¬C and have minimum impact on other rules; this increases the denominator of Expression (12) while keeping the numerator and Con$f$(r:B→C), Con$f$(rb2:D;B→A), and Con$f$(r: D;B →C) unaltered.

There is another way to provide indirect rule protection. Let us rewrite Inequality (10) as Inequality (13), where the confidences' of rb1 and rb2 rules are not constant conf (B→C)

$$> \frac{\dfrac{\text{Con}f\ (\text{rь1})}{\text{Con}f\ (\text{rь2})}\ (\text{con}\ f\ (\text{rь2}) + \text{con}f(\text{r:D,B}\rightarrow\text{C})\text{-1})}{\alpha} \tag{13}$$

Clearly, in this case Inequality (9) can be satisfied by increasing the Confidence of the base rule (B→C) of the redlining rule r: D;B→C to values greater than the right-hand side of Inequality (13) without affecting either the confidence of the redlining rule or the Confidence of the rb1 and rb2 rules. A possible solution for increasing Expression (8) in Inequality (13) to the target value is to perturb the class item from ¬C to C in the subset DBc of all records of the original data set which completely support the rule ¬A;B; ¬D→C and have minimum impact on other rules; this increases the numerator of Expression (8) while keeping the denominator and Con$f$(rb1 : A;B→D), Con$f$(rb2 : D;B →A), and Con$f$(r : D;B→C) unaltered.

6. Conclusion

In sociology, discrimination is the prejudicial treatment of an individual based on their membership in a certain group or category. It involves denying to members of one group opportunities that are available to other groups. Like privacy, discrimination could have negative social impact on acceptance and dissemination of data mining technology. Discrimination is a very important issue when considering the legal and ethical aspect. In order to prevent both direct and indirect discrimination in a dataset, a first step consists in discovering whether there exists direct or indirect discrimination. If any discrimination is found, the dataset is modified until discrimination is brought below a certain threshold or entirely eliminated.

7. References

[1] Sara Hajion & josep Domingo-Ferrer, Fellow "A Methodology for Direct and Indirect Discrimination Prevention in Data Mining," Knowledge and Data Engineering

[2]T. Calders and S. Verwer, "Three Naive Bayes Approaches for Discrimination-Free Classification," Data Mining and Knowledge Discovery, vol. 21, no. 2, pp. 277-292, 2010.

[3] European Commission, "EU Directive 2004/113/EC on Antidiscrimination," http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:373:0037:0043:EN:PDF, 2004.

[4] European Commission, "EU Directive 2006/54/EC on Antidiscrimination," http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:204:0023:0036:en:PDF, 2006.

[5] S. Hajian, J. Domingo-Ferrer, and A. Martı´nez-Balleste´,"Discrimination Prevention in Data Mining for Intrusion and Crime Detection," Proc. IEEE Symp. Computational Intelligence in Cyber Security (CICS'11), pp-47-54, 2011.

[6] S. Hajian, J. Domingo-Ferrer, and A. Martı´nez-Balleste´, "Rule Protection for Indirect Discrimination Prevention in DataMining," Proc. Eighth Int'l Conf. Modeling Decisions for Artificial Intelligence (MDAI '11), pp. 211-222, 2011.

[7] F. Kamiran and T. Calders, "Classification without Discrimination," Proc. IEEE Second Int'l Conf. Computer, Control and Comm.(IC4 '09), 2009.

[8] F. Kamiran and T. Calders, "Classification with no Discrimination by Preferential Sampling," Proc. 19th Machine Learning Conf.Belgium and The Netherlands, 2010.

[9] F. Kamiran,T. Calders, and M. Pechenizkiy, "Discrimination Aware Decision Tree Learning," Proc. IEEE Int'l Conf. Data Mining (ICDM '10), pp. 869-874, 2010.

[10] R. Kohavi and B. Becker, "UCI Repository of MachineLearningDatabases,"http://archive.ics.uci.edu/ml/datasets/Adult, 1996.

[11] D.J. Newman, S. Hettich, C.L. Blake, and C.J. Merz, "UCIRepository of Machine Learning Databases," http://archive.ics.uci.edu/ml, 1998.

[12] D. Pedreschi, S. Ruggieri, and F. Turini, "Discrimination-Aware Data Mining," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (KDD '08), pp. 560-568, 2008.

[13] D. Pedreschi, S. Ruggieri, and F. Turini, "Measuring Discrimination in Socially-Sensitive Decision Records," Proc. Ninth SIAM Data Mining Conf. (SDM '09), pp. 581-592, 2009.

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

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

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

[17] P.N. Tan, M. Steinbach, and V. Kumar, Introduction to DataMining. Addision-Wesley 2006.