A SURVEY ON DISCRIMINATION DETERRENCE IN DATA MINING

Arati Shrimant Mote[1], G. P. Chakote[2]

[1]B.A.M.U. University, Matsyodari Shikshan Sanstha's College of Engineering and Technology, Jalna, Maharashtra, India
[2]B.A.M.U. University, Matsyodari Shikshan Sanstha's College of Engineering and Technology, Jalna, Maharashtra, India

Abstract

For extracting useful knowledge which is hidden in large set of data, Data mining is a very important technology. There are some negative perceptions about data mining. This perception may contain unfairly treating people who belongs to some specific group. Classification rule mining technique has covered the way for making automatic decisions like loan granting/denial and insurance premium computation etc. These are automated data collection and data mining techniques. According to discrimination attributes if training data sets are biases then discriminatory decisions may ensue. Thus in data mining antidiscrimination techniques with discrimination discovery and prevention are included. It can be direct or indirect. . When choices are created depending on delicate features that period the discrimination is oblique. The elegance is oblique when choices are created depending on nonsensitive features which are strongly correlated with one-sided delicate ones. The suggested system tries to deal with elegance protection in information exploration. It suggests new improved techniques applicable for immediate or oblique elegance protection independently or both simultaneously. Conversations about how to clean coaching information sets and contracted information places in such a way that immediate and/or oblique discriminatory decision guidelines are transformed to genuine classification guidelines are done. New analytics to evaluate the utility of the suggested methods are suggests and comparison of these methods is also done.

Keywords: Antidiscrimination, information exploration, oblique and immediate elegance protection, concept protection, concept generalization, privacy.

--***---

1. INTRODUCTION

In sociology, elegance is considered as the harmful treatment of an individual depending on their account in a certain team or classification [1]. includes declining to the members of one team opportunities which are available to other groups. There is a list of an antidiscrimination functions. These laws are designed to limit elegance on the basis of a number of features in different configurations. The basic features can be race, religion gender, nationality, impairment, marriage position, and age. Different configurations can be employment and coaching, access to public services, credit and insurance. Even if there are some guidelines against elegance, all are sensitive, not practical. Technology can add proactively to regulation by contributing elegance finding and protection methods. Discrimination is nothing but counterproductive treatment of an individual. It is based on their membership in the certain group of their membership in a certain group or categories. It absorbs contradicting to members of one group opportunities. These opportunities are those which are available to another group. To avoid the discrimination, some laws are intended.

For large amount of data automatic and routine collection is allowed by the information society. In view of making automated decisions, this data is used to train association/classification rules. Automated decisions like loan granting /denial, top quality calculations, employees

choice, etc A feeling of equity is given by automating choices at first sight: category guidelines do not information themselves by selections. But in closer look, it is noticed that the category guidelines are actually discovered by the program from the coaching information. The discovered design can show a discriminatory prejudiced actions if the coaching information are one-sided for or against a particular group. That is nothing but program may conclude that purpose for loan refusal is just a being international. It is extremely suitable that the finding prospective tendencies and removing them from the training information without damaging their choice application. As the information exploration projects producing discriminatory designs from one-sided information places as part of the computerized choice, we must avoid information exploration from becoming itself a resource of elegance. It is experimentally shown in that the information exploration can be both a resource of elegance and a means for finding elegance.

2. LITERATURE SURVEY

Pedreschi et al proposed finding of discriminatory choices at very first. This was depending on exploration category guidelines (the inductive part) and thinking on them (the deductive part). It uses quantitative actions of elegance that formalize lawful explanations of elegance. Consider example of US Equal Pay Act. It says that a choice amount for any competition, sex, or cultural team which is less than four-fifths of the amount for the team with the highest

amount will usually be considered as proof of negative effect. This strategy is innovative to include mathematical importance of the extracted styles of elegance in [4] and to purpose about positive action and favoritism [5]. It is applied as an Oracle-based tool

Presently available elegance finding techniques consider each concept independently. Each concept is used for calculating elegance without considering other guidelines or the regards between them. Discrimination protection is the other big antidiscrimination objective in information exploration. It consists of causing styles which do not cause to discriminatory choices even if the unique coaching information places are one-sided. There are three approaches Preprocessing: Convert the resource information in such a way that the discriminatory tendencies in the unique information are eliminated so that no decision concept can be excavated from the modified information and implement any of the conventional information exploration techniques. The preprocessing techniques of data transformation and hierarchy-based generalization can be tailored from the comfort maintenance literary works. Along this line, [7], [8] perform controlled distortions of the coaching information from which a classifier is discovered by creating non-invasive variations resulting in an impartial data set. The preprocessing strategy is useful for programs in which a information set should be released and/or in which information exploration needs to be performed also by exterior events (and not just by the information holder).In-processing: The information exploration criteria is modified in such way that the unjust decision rule does not included by the causing designs. For example, for washing the elegance from the unique information set is suggested in which the nondiscriminatory limitation is included into a choice shrub student by changing its gashing requirements and trimming techniques during a novel foliage relabeling techniques.

1. Though, it is accessible that in-processing discrimination avoidance techniques must wait on new particular purpose data mining algorithms.
2. Post-processing: Rather than cleaning the genuine data set or altering the data mining algorithm process of post-processing modified the resulting data mining models. For example in [10], a confidence altering approached is proposed for the classification rule inferred by the CPAR algorithm. The authority to publishing the data is not given by the post-processing. They published only the modified data mining models. Since the process of data mining can be performed by data holder only.

In Classification with no Discrimination by Preferential Sampling [8] categorization without intolerance by Preferential Sampling is an expected solution for the discrimination like issues. It gives guaranteeing results with both steady and shaky classifiers. It decreases the security level by keeping up a high precision level. It gives comparable execution to "massaging" however without changing the dataset and dependably beats the "reweighing" plan.

In Integrating Induction and Assumption for judgment Evidence of Discrimination on [5] presented a reference model for the examination and revelation of discrimination in socially-sensitive choices taken by DSS. The methodology comprises first of extracting frequent classification rules, and afterward of examining them on the premise of quantitative measures of discrimination and their measurable significance. The key legitimate ideas of protected-by-law groups, direct discrimination, indirect discrimination, honest to goodness occupational prerequisite, affirmative activities and partiality are formalized as explanations over the set of concentrated runs and, perhaps, extra foundation information.

In Data Mining for Discrimination Discovery by Ruggieri et al. (2010) [3] present the issue of finding discrimination through data mining in a dataset of recorded choice records, taken by people or via programmed frameworks. They formalize the techniques of direct and indirect discrimination revelation by displaying protected-by-law groups and connections where discrimination happens in a classification based extraction. Essentially, classification rules extracted from the dataset permit for divulging connections of unlawful discrimination, where the level of load over protected-by-law groups is formalized by an augmentation of the lift measure of a classification rules.

In DCUBE: Discrimination Discovery in Databases by Turini et al. (2010) [6] says that DCUBE is a tool which is supporting the interactive and iterative process of detection of discrimination. The future users of DCUBE include: antidiscrimination establishment, proprietors of socially susceptible decision databases, and researchers, auditors in social sciences, economics and law.

A Survey regarding Association Rule Hiding technologies for Privacy in [11] presents classification and a survey of recent approaches that have been practical to the association rule hiding difficulty. Association rule hiding refers to the process of acclimatizing the original database in a way that influenced responsive alliance rules vanish without gravely affecting the data and the non-sensitive rules.

In Rule Protection for not direct Discrimination avoidance in information Mining in (2011) [10] displayed the first technique for avoiding not direct discrimination in information mining because of biased training datasets. There involvement in this paper centered on delivering training data which are free or about free from oblique discrimination while safeguarding their helpfulness to data mining algorithms. So as to avoid indirect discrimination in a dataset, a first step comprises in discovering whether there exists indirect discrimination. If any discrimination is originate, the dataset is changed until discrimination is brought beneath a certain threshold or completely removed.

Discrimination Prevention in Data Mining for Intrusion and Crime Detection by Domingo-Ferrer et al. (2011) [12] analyzed how discrimination could affect on cyber security applications, particularly IDSs. IDSs use computational

knowledge advances, for example, data mining. It is evident that the training data of these frameworks could be discriminatory, which would bring about them to settle on discriminatory decision when foreseeing interruption or, all the more for the most part, wrongdoing.

From [13] studied three Bayesian methods for discrimination-aware classification. They are altering the experiential probabilities in a Naïve Bayes model in such a way that its forecasts become discrimination-free. Next method concerned learning two dissimilar models; they are S_0 and S_1, and complementary these models afterwards. The last and most concerned method they initiated a latent variable *L* reflecting the latent "true" class of an object without discrimination.

In [1] exhibited two new algorithms, Apriori and Aprioritid, for finding all huge association rules between items in a vast database of transactions. We contrasted these algorithms with the awhile ago known algorithms, the AIS [4] and SETM [13] algorithms.

As concluded in EU Directive 2004/113/EC on Anti-Discrimination (2004) [14] there might be another attributes for example zip code which are highly associated with the sensitive ones and permit inferring discriminatory rules. Since the two most significant challenges concerning the discrimination prevention is; first challenge is instead of considering only direct discrimination we want to consider both direct and indirect discrimination. Second challenge is to discovering the good arrangement among the removes of discrimination and the excellence of the resultant training data sets and the model of information mining.

In [7] suggest the thought of discrimination is non minor and poses ethical furthermore legitimate issues and also hindrances in commonsense applications. CND furnishes us with a basic yet influential beginning stage for the arrangement of the discrimination issue. CND classifies the future information (both discriminatory and non discriminatory) with least discrimination and high exactness. It likewise addresses the issue of redlining.

Since some methods are already proposed for each approach we mention above, the concept of prevention of discrimination is the topic of research for the researchers.

3. BACKGROUND

In this section we will discuss about the background knowledge essential for this paper. Initially we will remember some basic definition related with the data mining. Later on we convoluted on measuring and determining discrimination.

4. BASIC DEFINITIONS

- Data set is a gathering of data objects and their qualities Suppose DB is the original database.
- An item is characteristic along with its value, for example Nationality – Indian.

- An item set Y is collect than of more than items, for example {Nationality – Indian, State - Gujarat}
- Classification rule is an expression Y→ C, where C is a class item and Y is an item set.
- Support of an item set, supp(Y), it is the division of records which contains the item sets Y. we can say that rule Y→ C is completely sustain by the record if Y and C both are present in the record.
- Confidence of a classification rule, conf (Y→ C) evaluates the how frequently the class item C appears in record which contains Y. Therefore, if supp (Y) > 0 then,

$$conf\,(Y \rightarrow C) = \frac{supp(Y,C)}{supp(Y)} \qquad (1)$$

- Common categorization rule is a classification rule with the sustain and confidence value which is greater than the specific lower bounds. In which support is the measure of statistical significance and confidence is the measure of strength of rule.

5. PD AND PND CLASSIFICATION RULE

Suppose DC be the set of prearranged prejudiced items in the DB. FP is the frequent classification rule which fall into the following classes:

1. Classification rule Y→ C is PD when Y = A, B with A is the subset of DC a nonempty discriminatory item set and B is the nondiscriminatory item set.
2. Classification rule Y→ C is PND where Y = D, B is a nondiscriminatory item set.

The phrase 'potentially' means a PD rule probably lead to discriminatory decision. As a result, for quantifying the direct discrimination potential some measures are needed.

6. PROPOSED SYSTEM

In this section we discussed about our proposed approached including data transformation methods which used for prevent of direct or indirect discrimination.

6.1 Proposed Approach

The proposed approach for prevention of direct and indirect discrimination is split up into two phases:

- **Measurement of discrimination**
  Direct and indirect discrimination detection contains obtaining the alpha discriminatory rules and redlining rules. Initially, PD and PND rules are generated on the basis of discriminatory items in the database DB and FP the frequent classification rule. After that by using the direct discriminatory measures and the discriminatory threshold the direct discrimination is measured by obtaining the alpha discriminatory rules with the PD rules. After that, same as the direct discrimination, indirect discrimination is measured by obtaining the redlining rules with the PND rules combining with

the background knowledge, using an indirect discrimination measures and the discrimination threshold.

- **Transformation of discrimination**
  Transforming the original database DB in such way that to eliminate direct or indirect discriminatory, together with least impact on the data and on rightful decision rules, so that no inequitable decision rule can be excavate from the transaction database
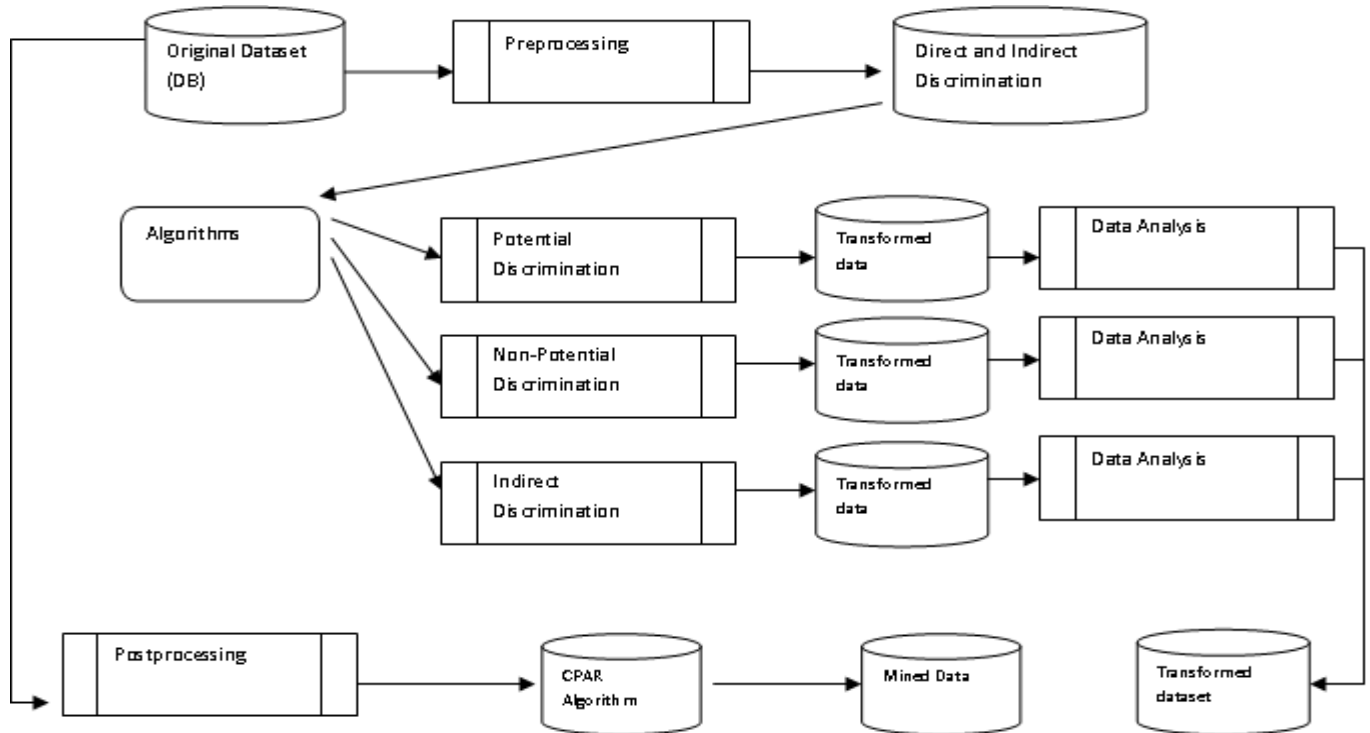
7. COMPARATIVE ANALYSIS:

Here comparative study to understand how different methods affects the discrimination in mining. New different researches go under for finding reliable result with no discrimination present in databases without data loss.

Table 1: Different Mining Techniques for Discrimination with Advantages and Disadvantages

Paper References	Methods /Techniques used	Advantages	Disadvantages
"Fast Algorithms for Mining Association Rules in Large Databases", in 1994 [1]	- Apriori, AprioriTid are designed - comparison between AIS and SETM algorithms with Apriori, AprioriTid	- Association rule method for the large database. - get large item sets to work.	- Performance gap is increases with the problem size increasing.
"Classification with no Discrimination by Preferential Sampling", 2010 [8]	- it's for less intrusive technique to get unbiased data without changing class labels as efficient as result of massaging and reweighing by sampling algorithm.	- introduced the idea of Classification with No Discrimination (CND).	- did not consider such as do not proposing discrimination model which is used in many cases.
"Rule Protection for Indirect Discrimination Prevention in Data Mining", in 2011 [10]	- study to find combinations of attributes to get indirect discrimination with preprocessing with algorithm- 1,2 [10] - Prevention degree and protection preservation (DPD and DPP) used for results	- antidiscrimination in the context of cyber security. - proposed data transformation method for indiscrimination prevention and considered several discriminatory attributes and their combinations.	- does not run method on real dataset. - don't consider background knowledge.

"Three Naive Bayes Approaches for Discrimination-Free Classification", in 2010 [13]	- Naive Bayes approach with 3 methods (1)change probability like no discrimination in predictions (2) balance models with S = 0/ 1 (3) true class with no discrimination by Latent variable.	- presented a modified Naive Bayes classification approach for taking decisions easily.	- doesn't consider numerical attributes viz. Income as a sensitive attribute.
" EU Directive 2006/54/EC on Anti-Discrimination" in 2006[14]	- Social information case discrimination and for that antidiscrimination study.	- Presented the issue of discrimination in social sense viz. against minorities and disadvantaged groups.	- Not focused on the rule based framework process for direct and indirect discrimination.

8. CONCLUSION

As discrimination is a very important issue of data mining. The purpose of this paper was to develop new preprocessing discrimination prevention including different data transformation methods that can prevent direct discrimination, indirect discrimination along with both at the same time. Also it consists of supporting discrimination discovery, namely the unveiling of discriminatory decisions hidden, either directly or indirectly, in a dataset of historical decision records, possibly built as the result of applying a classifier. As future work we are exploring measures of discrimination different from the ones considered in this paper along with privacy preservation in data mining. Further, implementation of discrimination prevention in Post-processing will be done. Proposed Algorithm achieves high accuracy and efficiency.

REFERENCES

[1] Sara Hajian and Josep Domingo-Ferrer," A Methodology for Direct and Indirect Discrimination Prevention in Data Mining", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 7, JULY 2013.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

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

[11] V. Verykios and A. Gkoulalas-Divanis, "A Survey of Association Rule Hiding Methods for Privacy," Privacy-Preserving Data Mining: Models and Algorithms, C.C. Aggarwal and P.S. Yu, eds., Springer, 2008.

[12] 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.

[13] 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.

[14] European Commission, "EU Directive 2004/113/EC on Anti-Discrimination," http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:373:0037:0043:EN:PDF, 2004.

BIOGRAPHIES

Arati Shrimant Mote received the B.E. degree in Computer Science and Engineering from Marathwada Institute of Technology in 2010. Now doing PG in Computer Science and Engineering from Matsyodari Shikshan Sanstha's College of Engineering and Technology, Jalna, Maharashtra, India

Prof. G. P. Chakote is working as Professor at Matsyodari Shikshan Sanstha's College of Engineering and Technology, Jalna, Maharashtra, India