

Prevention of Discrimination in Data Mining

Miss. Melanie Ann Thomas, Mrs. Joshila Grace (lecturer)

Sathyabama University

Abstract: *Data mining is the process of analyzing data from different perspectives and summarizing it into useful information. Automated data collection allows users to analyze data from many different dimensions or angles, categorize it and summarize the relationships identified. Discrimination is of two types, Direct and Indirect discrimination. Direct discrimination occurs when decisions are made based on sensitive attributes. Indirect discrimination occurs when decisions are made on non sensitive attributes which are strongly related with sensitive. In this paper the system tackles discrimination in data mining by using methods like Rule protection, Rule generalization and eligibility of the client is done by using the Preferential algorithm. This is applicable for both direct and indirect discrimination. The cleaning training data sets in such a way that direct and/or indirect discriminatory decision rules are converted to nondiscriminatory classification rules is discussed in this paper.*

Index Terms: *Antidiscrimination, data mining, direct and indirect discrimination prevention, Rule protection, Rule generalization, Preferential sampling.*

I. Introduction-

The word discrimination originates from the Latin word *dis-criminare*, which means to “distinguish between”. In social sense, however, discrimination refers specifically to an action based on prejudice resulting in unfair treatment of people on the basis of their membership to a category, without regard to individual merit. As an example, U.S. federal laws[4] prohibit discrimination on the basis of race, color, religion, nationality, sex, marital status and age in a number of settings, including: credit/insurance scoring(Equal Credit Opportunity Act); sale, rental, and financing of housing (fair Housing Act).

Concerning the research side, the issue of discrimination in credit, mortgage, insurance, labor market, education and other human activities has attracted much interest of researchers in economics and human science.

Data mining can be both a source of discrimination and a means for discovering discrimination [2]. Direct discrimination is evident when a person is treated unfavorable because of his or her personal attribute like sex, race, age, disability or parental status. This is straightforward and can affect the person being discriminated seriously. Indirect discrimination occurs when a certain policy appears to deal with all people equally but has the result of affecting a number of certain people where the requirement is unreasonable. The policy may seem harmless, but it has a discriminatory result against certain individuals.

II. Existing System-

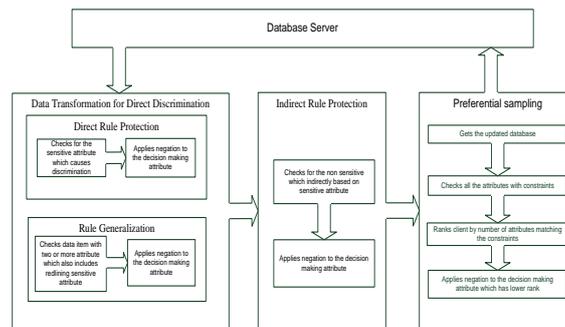
A discrimination prevention technique consists of patterns that do not lead to discriminatory decisions even if the original data sets are biased. Three approaches possible are preprocessing, in-processing and post-processing. This system concentrates on discrimination prevention based on preprocessing, because the preprocessing approach seems the most flexible one it does not require changing the standard data mining algorithms, unlike the in-processing approach, and it allows data publishing (rather than just knowledge publishing), unlike the post-processing approach. Collection of large amounts of data to be trained by classification rules in view of making automated decisions, like loan granting/denial, insurance premium computation, personnel selection, etc. Classification rules are achieved by training a large database based on the certain rules like gender, area, religion, etc. thus it exposes the direct discrimination occurred on which the discrimination prevention is made on the original data set. Whereas this system have some limitation, thus paved the way to propose new discrimination prevention technique called rule protection and rule generalization to give protection against both direct and indirect discrimination.

III. Proposed system-

Preprocessing methods are proposed to overcome the limitations in existing system. The data transformation methods (Rule protection and rule generalization) are based on measures of both Direct and Indirect discrimination and also can deal with several discriminatory items. The system provides utility measure and its approach to discrimination prevention is broader here.

A database with large dataset is designed which includes both direct and indirect discriminatory item sets. The database is maintained by a large database server by which the various discrimination removal methods are applied. These methods (direct rule protection, rule generalization and indirect rule protection) are employed on the main database with the discrimination. After applying discrimination methods, discrimination measurement is done, and then to the updated database PREFERENTIAL SAMPLING algorithm is used. Checks all attributes with constraints, then ranks client by number of attributes matching constraints, after that applies negation to the decision making attribute which has lower rank [3].

IV. Architecture Diagram



V. Methodology-

5.1 System Model- In this module the designing of database with large data sets is done and it has complied as the database with both direct discriminatory and non discriminatory items. The large database is maintained by database server to which all discrimination removal methods are applied.

5.2 Discrimination Measurement- Discriminatory rules and indirect rules are used to identify Direct and Indirect discrimination. The system divides the classification into two groups: PD (Potential Discriminatory) and PND (Potential Non Discriminatory) rules based on predetermined discriminatory items in database. By identifying discriminatory rules among Potential Discriminatory rules using a direct discrimination measure is the way how system measures direct discrimination. The system measures the indirect discrimination by identifying indirect rules among the Potential Non Discriminatory rules. The system takes the database of direct and indirect discriminatory rules obtained by above process.

5.3 Data Transformation for Direct Discrimination

5.3.1 Direct Rule Protection- This module takes the database of direct discrimination. It converts each discriminatory rule into a protective rule. The system extracts the main discriminatory item sets causing the direct discrimination, then takes each indirect rule from this item sets. Then negation of decision making attribute is done. At last a database is generated by removing all discrimination causing item sets.

5.3.2 Rule Generalization- The system considers relation between rules instead of discrimination measure. To formalize the dependency among rules, the system takes combined set of general rule and redlining rule and performs the negation process on the decision making attribute.

5.3.3 Direct Rule Protection and Rule Generalization- Here at first direct rule protection and then rule generalization is done, then it generates a database which will be from direct discrimination.

5.4 Data Transformation for Indirect Transformation

5.4.1 Indirect Rule Protection- Using the above specified process the system extracts all the item sets from the database, which can still get discriminated by the item set which are not directly resulted in discrimination.

5.5 Data Transformation for both Direct and Indirect Discrimination- This module simultaneously performs the both data transformation methods in direct and indirect discrimination. Then it generates the database which is free from both the direct discrimination and indirect discrimination occurring in the database.

5.6 Discrimination Measurement after applying Data Transformation methods- This module gets the final database which it got from data transformation methods in both direct and indirect discrimination. The next step is that it measures the discrimination data removed by the above methods and compares with the initial database of direct and indirect discriminatory rules by which it gives the efficiency of the Rule Protection and Rule Generalization.

5.7 Preferential Sampling Algorithm-

Introducing of Preferential Sampling (PS) scheme is to make the dataset bias free. Instead, PS changes the distribution of different data objects for a given data to make it discrimination free. The idea is that the data objects close to the decision boundaries are more prone to be the victim of discrimination. So the main objective is to change the distribution of these borderline objects to make the dataset discrimination free. A ranking function is used on the original data, to identify the datasets close to be discriminated.

STEPS:

1. Checks for eligibility criteria.
2. Ranks the client based on number of eligibility criteria they satisfy.
3. Then negation is applied (approve to deny or deny to approve) to those who having lower rank.

VI. Comparison between Existing and Proposed

Existing System	Proposed System
<ol style="list-style-type: none"> 1. Existing system attempts to detect discrimination in the original data only for one discriminatory item based on only one measure. 2. Existing System considered only direct discrimination. 3. Existing System cannot guarantee that the transformed data is discrimination free. 4. Existing System does not include any measure, to evaluate how much of the discrimination is removed. 5. Existing System does not include any metrics to find how much information was lost. 	<ol style="list-style-type: none"> 1. Proposed System attempts to detect discrimination in the original data not only for one discriminatory item but also for all discriminatory items occurred. 2. Proposed System considers both direct and indirect discrimination 3. Proposed System guarantees that the transformed data set is really discrimination free. 4. Proposed System includes measure to evaluate how much of discrimination is removed. 5. Proposed system include metrics to find out how much information was lost.

VII. Conclusion-

The usage of preferential sampling after applying all of the discrimination measure is an excellent solution to the discrimination problem. It gives promising results with both stable and unstable classifiers. It reduces the discrimination level by maintaining a high accuracy level and information loss is very less.

References-

- [1]. Sara Hajian and Josep Domingo "A methodology for Direct and Indirect Discrimination Prevention in Data Mining" IEEE Transaction on Knowledge and Data Engineering Vol.25. No.7. July 2013.
- [2]. D.Pedreschi, S.Ruggieri and F.Turini, " Discrimination-Aware Data Mining," Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining.
- [3]. Faisal Kamiran, Toon Calders "Classification with no Discrimination by Preferential Sampling" Eindhoven University of Technology, The Netherlands.
- [4]. Dino Pedreschi Salvatore Ruggieri Franco Turini Discrimination-aware Data Mining Dipartimento di Informatica, Università di Pisa L.go B. Pontecorvo 3, 56127 Pisa, Italy