ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

Survey on Discrimination Techniques for Privacy Preserving Data Mining

S.Subbulakshmi[1], B.Arulkumar [2], Syed Farmhan.S[3]

PG Scholar, Computer Science and Engineering, Vivekanandha College of Engineering for Women, Namakkal, India[1]

Assistant Professor, Computer Science and Engineering, Vivekanandha College of Engg., for Women, Namakkal, India[2]

Assistant Professor, Computer Science and Engineering, Vivekanandha College of Engg., for Women, Namakkal, India[3]

Abstract: Data mining is an important technology for extracting useful knowledge hidden in large collections of data. Privacy is a main issue in Data mining. The former is an unintentional or planned admission of a user profile or activity data as part of the output of a data mining algorithm or as a result of data sharing. For this reason, privacy preserving data mining has been introduced to trade-off the utility of the resulting data for protecting individual privacy. Along with privacy, discrimination is a very important issue when considering the legal and ethical aspects of data mining. Automated data collection and data mining techniques such as classification rule mining have covered the way to making automated decisions, like loan granting, insurance premium computation, etc. If the training data sets are unfair in what regards discriminatory (sensitive) attributes like age, race, gender, religion, etc., discriminatory decisions may proceed. For this basis, antidiscrimination techniques including discrimination discovery and prevention have been introduced in data mining. To solve such problems there are some algorithms presented by various authors worldwide. The primary goal of this survey paper is to understand the existing prevention techniques and to achieve efficiency.

Keywords: Data Mining, Discrimination, Privacy Preserving, Decision Tree, Rules.

I. INTRODUCTION

Data mining and knowledge discovery in databases are two new research areas that investigate the automatic extraction of previously unknown patterns from large amounts of data. Data mining techniques are used in business and research and are becoming more and more popular with time. Data mining involves the extraction of implicit previously unknown and potentially useful knowledge from large databases. Data mining is a very challenging task since it involves building and using software that will manage, explore, summarize, model, analyses and interpret large datasets in order to identify patterns abnormalities. The important issue of data mining is privacy. Privacy preserving in data mining techniques are being used increasingly in wide verity of application. Privacy Preserving Data Mining (PPDM) is a research area concerned with the privacy driven from personally identifiable information when considered for data mining. Therefore, PPDM has become an increasingly important field of research. PPDM is a novel research direction in data mining.

Along with privacy, discrimination is a very important issue when considering the legal and ethical aspects of data mining. Discrimination can be viewed as the act of illegally treating people on the basis of their belonging to a specific group. For instance, individuals may be discriminated because of their race, ideology, gender, etc.especially when those attributes are used for making decisions about them like giving them a job, loan, insurance, finance, etc. Discovering such potential biases and eliminating them from the training data without harming their decision-making utility is therefore highly desirable. For this basis, Anti-discrimination techniques including discrimination discovery and prevention have been introduced in data mining.

Discrimination can be either direct or indirect (also called systematic).Direct discrimination consists of set of laws(rules) or procedures(events) that explicitly mention minority or deprived groups based on sensitive discriminatory attributes related to group membership. Indirect discrimination consists of set of laws (rules) or procedures that, while not clearly mentioning discriminatory attributes, deliberately or not deliberately could generate discriminatory decisions. Redlining by financial institutions (refusing to grant mortgages or insurances in urban areas they consider as deteriorating) is an archetypal example of indirect discrimination, although definitely not the only one. With a slight neglect of language for the sake of compression, in this paper indirect discrimination will also be referred to as redlining and rules causing indirect discrimination will be called redlining rules [7]. Indirect

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

discrimination could happen because of the availability of some background knowledge (rules) [1], because of the existence of nondiscriminatory attributes that are highly correlated with the sensitive ones in the original data set.
The main charity of this paper is to provide the best solution for removing direct and/or indirect discrimination biases in the original data set while preserving data quality.

II. Literature Survey

There are some methods presented by various authors worldwide. Those methods can be analysed and provide a limitations are given below.

A. Two naive Bayes model

This approach is to avoid this dependence on A_s by removing the correlation between S and A_s from the data-set used to train the naive Bayes classifier. This can for occurrence be achieved by removing as from the data-set altogether; the resulting classifier will be independent without modification. The model M+ is learned using only the tuples from the data-set that have a favored sensitive value S+. The model M− uses only those that have a discriminated sensitive value S−. The overall classifier chooses either M+ or M− depending on the value of S and uses that model's classification Overall, since M+ and M− share the same naive Bayes structure, this approach can be modeled by connecting S to all other attributes in this structure. This is shown by the graph in Fig.1. Since all possibility distributions in the naive Bayes structure depend on S, this equals two different naive Bayes models. In this overall model, we eliminate discrimination by modifying the probability P (C|S).

$$P(C, S, A1..., A_n) = P(C)\ P(S|C)\ P(A_1|C) \ldots P(A_n|C)$$

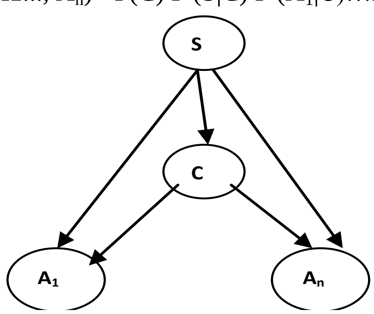

Fig 1 Two Naive Bayes model

1) Result:

The 2 naive Bayes models method has the lowest dependence on S, resulting in only about 5% discrimination if Sis removed. This is somewhat surprising since this model uses S to split the data and then learn two separate models. It appears that, these two separate models are good at estimating S from the other attributes $A_1,\ldots,A_n$. This method performs best: it achieves high accuracy scores with zero discrimination, and has the smallest dependency on S.

2) Drawbacks:

The main drawback of this approach is not applicable for Indirect discrimination and the accuracy of data could be low,it cannot measure the utility rate of discrimination from the original data set.

B. Preferential Sampling

Introduced the idea of Classification with No Discrimination (CND). We propose a new solution to the CND problem by we introduce a Preferential Sampling (PS) scheme to make the dataset bias free. Instead, PS changes the distribution of different data objects for a given data to make it discrimination free. To identify the borderline objects [5], PS starts by learning a ranker on the training data. PS uses this ranker to class the data objects of DP and PP in ascending order, and the objects of DN and PN in descending order; both w.r.t. the positive class probability. Such understanding of data objects makes sure that the higher the rank an element occupies, the closer it is to the borderline.PS starts from the original training dataset and iteratively duplicates (for the groups DP and PN) and removes objects (for the groups DN and PP) in the following way:
Decreasing the size of a group is always done by removing the data objects closest to the borderline. Increasing the sample size is done by duplication of the data object closest to the borderline.
Figure 2 gives an illustration of Preferential Sampling (PS), showing 40 data point. Data points of the desired class and the negative class are represented by + and – symbols respectively [5].
PS works in the following steps:
(i) Divide the data objects into the four groups, DP, DN, PP, and PN.
(ii) Any ranking algorithm may be used for calculating the class probability of each data tuple. This ranking will be used to identify the borderline data objects.
(iii) Calculate the expected size for each group to make the dataset bias free.
(iv) Finally apply sampling with replacement to increase the size of DP and PN. And decrease the size of DN and PP.

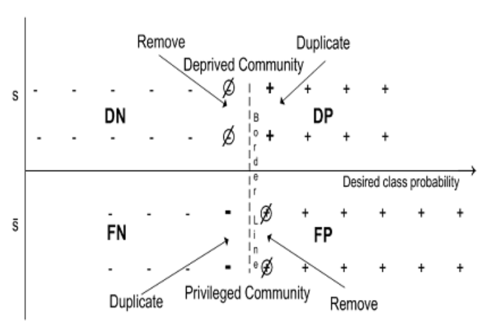

Fig 2: Pictorial representation of Preferential Sampling scheme. The re substituted data points are in bold while the encircled ones are skipped.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

1) Result:
Classification with No Discrimination by Preferential Sampling is an excellent solution to the discrimination problem. It gives promising results with both stable and unstable classifiers give more accurate results but do not reduce the discrimination

2) Drawbacks:
Low data utility rate and minimum discrimination removal. This PS is also not applicable for Indirect discrimination.

C. Decision Tree Learning

This approach in which the non-discriminatory constraint is pushed deeply into a decision tree learner by changing its splitting criterion and pruning strategy by using a novel leaf relabeling approach. We propose the following two techniques for incorporating discrimination awareness into the decision tree construction process:

Dependency-Aware Tree Construction: When evaluating the splitting criterion for a tree node, not only its contribution to the accuracy, but also the level of dependency caused by this split is evaluated.

Leaf Relabeling: Normally, in a decision tree, the label of a leaf is determined by the majority class of the tuples that belong to this node in t he training set. In leaf relabeling we change the label of selected leaves in such a way that dependency is lowered with a minimal loss in accuracy. We show a relation between finding the optimal leaf relabeling and the combinatorial optimization problem KNAPSACK [12]. Based on this relation an algorithm [6] is proposed.

1) Result:
This method gives high accuracy and low discrimination scores when applied to non-discriminatory test data. In this scenario, our methods are the best choice, even if we are only concerned with accuracy. (2) The enrichment in discrimination reduction with the relabeling method is very satisfying. The relabeling reduce discrimination to almost 0 in almost all cases if we decrease the value of ^{2} to 0. (3) The relabeling methods out-perform the baseline in almost all cases. As such it is reasonable to say that the straightforward solution is not satisfactory and the use of dedicated discrimination-aware techniques is justified. (4) This methods significantly improve the current state-of-the-art techniques [3] w.r.t. accuracy discrimination trade off.

2) Drawbacks:
The result of this approach has mostly similar to the Naive Bayesian Approach and it only concerned with accuracy. Discrimination removal is very low using relabeling method.

D. Pre-processing Approach

The key contributions of this model are as follows: (1) proposing a new discrimination prevention method based on data transformation that can consider several discriminatory attributes and their combinations; (2) proposing some measures for evaluating the proposed method in terms of its success in discrimination prevention and its impact on data quality.

Data Transformation Method:
An appropriate data transformation method is required to modify original data in such a way that the transformation requirement for each α-discriminatory rule is satisfied without seriously affecting the data or the non α-discriminatory rules. Based on these objectives, the data transformation method should increase or decrease the confidence [5] of the rules to the target values with minimum impact on data quality, that is, maximize the disclosure prevention measures and minimize the information loss measures.

The transforming (db′) with minimum impact(db′) could reduce the impact of this transformation on turning the α-protective rules to α-discriminatory rules and on generating the extractable rules from original dataset in the transformed dataset.

Utility Measures:
The transformation method should be evaluated based on two aspects:

- The success of the above solution in removing all evidence of discrimination from the original dataset (degree of discrimination prevention).
- The contact of the proposed solution on data quality (degree of information loss).

The following measures are proposed for evaluating solution:

Discrimination Prevention Degree (DPD): This measure quantifies the percentage of α-discriminatory rules that are no longer α-discriminatory in the transformed dataset.

Discrimination Protection Preservation (DPP): This measure quantifies the percentage of the α-protective rules in the original dataset that remain α-protective rules in the transformed dataset.

Misses Cost (MC): This measure quantifies the percentage of rules among those extractable from the original dataset that cannot be extracted from the transformed dataset.

Ghost Cost (GC): This measure quantifies the percentage of the rules among those extractable from the transformed dataset that could not be extracted from the original dataset.

1) Result:
The DPD and DPP measures are used to evaluate the success of proposed method in discrimination prevention; ideally they should be 100%. The MC and GC measures are used for evaluating the degree of information loss (impact on data

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

quality); ideally they should be 0%. MC and GC were previously proposed as information loss measures for knowledge hiding in PPDM.

2) Drawbacks:

The measurement of discrimination has only deal with α-discriminatory rules. This approach is also not supported for indirect discrimination.

E. Indirect Discrimination Prevention

This Method regarding discrimination prevention is considering indirect discrimination other than direct discrimination and another challenge is to find an optimal trade-off between anti-discrimination and usefulness of the training data.

The main contributions of this method are as follows: (1) a new pre-processing method for indirect discrimination prevention based on data transformation that can consider several discriminatory attributes and their combinations; (2) some measures for evaluating the proposed method in terms of its success in discrimination prevention and its impact on data quality.

This solution is based on the fact that the dataset of decision rules would be free of indirect discrimination if it contained no redlining rule.

Data Transformation Method for Indirect Discrimination: Rule Protection

The indirect discriminatory measure (i.e. elb), to convert redlining rules into non-redlining rules, we should enforce the following inequality for each redlining rule r: D, B→C$_{in}$ RR:

$$elb\ (\gamma, \delta) < \alpha$$

In order to implement this data trans-formation method for indirect discrimination prevention, we simulate the availability of a large set of background rules under the assumption that the datasetcontains the discriminatory items. The utility measures of indirect discrimination is same as the above preprocessing approach based on the readlining rule dataset RR

1) Result:

The values of DDP and DPD achieves a high degree of indirect discrimination preventionin different cases (i.e. different values ofα). In addition, the values of MC and GC demonstrate that this proposed solution incurs little information loss, especially when α is not too small. By decreasing the value of α,the amount of redlining rules is increased, which causes further data transformation to be done, there by increasing MC and GC.

2) Drawbacks:

The execution time of this algorithm increases linearly with the number of redlining rules andα-discriminatoryrules.This method is only deal with indirect discrimination and it cannot measure the direct discriminatory items.

F. Direct and Indirect Discrimination Prevention Method

This new technique applicable for direct or indirect discrimination prevention individually or both at the same time and effective at removing direct and/or indirect discrimination biases in the original data set while preserving data quality.

This method can be described in terms of two phases:

Discrimination measurement- Direct and indirect discrimination discovery includes identifying α discriminatory rules and redlining rules.

(i) Based on predetermined discriminatory items in DB, frequent classification rules in FR are divided in two groups: PD and PND rules.

(ii) Direct discrimination is measured by identifying α-discriminatory rules among the PD rules using a direct discrimination measure (elift) and a discriminatory threshold (α).

(iii) Indirect discrimination is measured by identifying redlining rules among the PND rules combined with background knowledge, using an indirect discriminatory measure (elb), and a discriminatory threshold (α).

Data transformation- Transform the original data DB$_{in}$ such a way to remove direct and/or indirect discriminatory biases, with minimum impact on the data and on legitimate decision rules, so that no unfair decision rule can be mined from the transformed data.

Transformation Method:

The key problem of transforming data with minimum information loss to prevent at the same time both direct and indirect discrimination. We will give a pre-processing solution to simultaneous direct and indirect discrimination prevention.

There are two transformation method used in both direct and indirect discrimination removal.

(i) ***Direct Rule Production -*** In order to convert each α-discriminatory rule into a α-protective rule, based on the direct discriminatory measure.

$$elift\ (r') < \alpha$$

(ii) ***Indirect Rule Protection -*** In order to turn a redlining rule into an non-redlining rule, based on the indirect discriminatory measure (i.e., elb in Theorem 1)[6], we should enforce the following inequality for each redlining ruler: D,B→C$_{in}$ RR:

$$elb\ (\gamma, \delta) < \alpha$$

These two data transformation method for used simultaneous direct and indirect discrimination prevention.

Utility Measures:

These techniques should be evaluated based on two aspects.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

- To measure the success of the method in removing all evidence of direct and/or indirect discrimination from the original data set.
- To measure the impact of the method in terms of information loss

To measure discrimination removal, four metrics were used:

(i) ***Direct discrimination prevention degree (DDPD):*** This measure quantifies the percentage of α-discriminatory rules that are no longer α-discriminatory in the transformed data set.

(ii) ***Direct discrimination protection preservation (DDPP):*** This measure quantifies the percentage of the α-protective rules in the original data set that remain α-protective in the transformed data set.

(iii) ***Indirect discrimination prevention degree (IDPD):*** This measure quantifies the percentage of redlining rules that are no longer redlining in the transformed data set.

(iv) ***Indirect discrimination protection preservation (IDPP):*** This measure quantifies the percentage of non redlining rules in the original data set that remain non-redlining in the transformed data set.

The above measures are used to evaluate the success of the proposed method in direct and indirect discrimination prevention, ideally their value should be100 percent. The data quality is measured using the MC and GC.

1. Result:

This method achieves a high degree of both direct and indirect discrimination prevention for different values of the discriminatory threshold (α). The important point is that, by applying this method, we get good results for both direct and indirect discrimination prevention at the same time. In addition, the values of MC and GC demonstrate that the method incurs low information loss.

2. Drawbacks:

The main drawbacks of this method contain Low privacy assurance and Limited utility ratio of data. The association of privacy is not analysed from the transformed dataset.

III. PERFORMANCE ANALYSIS

This paper gives a comparison of state-of-the-art methods on the Census Income dataset. It affords the entitlement of discrimination and accuracy for the above discussed methods. The performances of various methods have been specified below Table 1 based on discrimination and accuracy.

TABLE I
THE RESULTS OVER THE CENSUS INCOME DATASET

Methods	Discrimination Removal (%)	Accuracy (%)
Two naive Bayes model	0.047	0.807
Preferential Sampling	0.17 ± 2.64	83.98 ± 1.12
Decision Tree Learning	10.95 ± 1.76	81:10±0 :47
Pre-processing Approach	70.6	1.96
Indirect Discrimination Prevention	90.90	1.62
Direct and Indirect Discrimination Prevention Method	98.8	0.69

IV. PROPOSED SOLUTION

The above drawback of direct and indirect discrimination prevention model is overcome by new techniques. The privacy is connection with current privacy models, like differential privacy. It will provide the high privacy rate. This method is integrated with the previous existing method of direct and indirect discrimination prevention mechanism and to find synergies between rule hiding for privacy-preserving data mining and association rule hiding for discrimination removal. Rule privacy is optimized with rule generalization mechanism. These methods provide the competent outcome of removing the discrimination with high privacy rate.

IV. CONCLUSION

In this paper, we have carried out a wide survey of the different approaches for discrimination prevention, and analyses the major algorithms available for discrimination prevention method and point out the drawbacks of direct and indirect discrimination prevention method. These above methods are evaluating over the census income dataset and the performances are shown in the TABLE I. Each method gave the better outcome of discrimination prevention. Our proposed solutions are overcome the above problems. Its only approximate to our goal of discrimination prevention, we need to further perfect those approaches or develop some efficient methods.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

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] 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.

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

[4] 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.

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

[6] 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.

[7] 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.

[8] 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. D. Pedreschi, S. Ruggieri, and F. Turini. Discrimination-aware data mining. In Proc. ACM SIGKDD, 2008.

[9] P.N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Addison-Wesley, 2006.

[10] 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.

[11] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Prosati. Complexity and Approximation. Combinatorial Optimization Problems and Their Approximability Properties. Springer, 2003.