International Journal of Advanced Technology in Engineering and Science
Vol. No.5, Issue No. 03, March 2017
www.ijates.com
ijates
ISSN 2348 - 7550

ANALYSIS ON DISCRIMINATION INTERRUPTION TECHNIQUES IN DATA MINING

D. Swetha[1], Pavani.G[2]

ABSTRACT

Extracting useful information hidden in large collection of data is known as data mining. Discrimination is a very important issue when considering the legal and ethical aspects of data mining. Discrimination means unfairly treating people on the basis of their cast, religion , gender etc. Discrimination can be either direct or indirect To solve this problem there are some algorithm presented by various authors worldwide. The main goal of this survey paper is to understand the existing prevention technique.

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 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.

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, Antidiscrimination techniques including discrimination discovery and prevention have been introduced in data mining.

Discrimination can be either direct or indirect. Direct discrimination consists of set of law or procedures that explicitly mention minority or deprived groups based on sensitive discriminatory attributes related to their membership on a specific group. Indirect discrimination consists of set of laws or procedures that, while not clearly mentioning discriminatory attributes, deliberately or not deliberately could generate discriminatory decisions. Redlining by financial institutions is an archetypal example of indirect discrimination.

II. LITERATURE SURVEY

There are different methods proposed to avoid such discriminatory behavior in data mining. Those methods can be analyzed and provide a limitations are given below.

ijates
ISSN 2348 - 7550

2.1 Two Naïve Bayes Model

In this method naive Bayes is modify for discrimination classification. Discrimination laws do not allow the use of these rules of attributes such as gender, religion. Using decision rules that base their decision on these attributes in classifier. The approaches are used in this paper Navies Bayes model, Latent variable model, and Modified naives Bayes. The naives bayes model is A bayes classifier is a simple possibility classifier based on applying bayes theorem with strong statistical independence assumption. Depending on precise nature of the probability model, naive bayes classifiers can be trained very efficiently in supervised learning. A latent variable model is a numerical model that relates a set of variables to set of latent variables. The responses on the indicators or manifest variables are the results of an individual's position on the latent variables. The modified naive bayes is Modify the probability distribution p(s/c) of the sensitive attribute values s given the class values.

2.1.1 Result

2 naive Bayes models method has the lowest dependence on S, resulting in only about 5% discrimination if S is 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 $A_1,...,A_n$. This method performs best: it achieves high accuracy scores with zero, and has the smallest dependency on S.

2.1.2 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.

2.2 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 , 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 with respect to 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 and removes objects 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.

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.

2.2.1 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.

2.2.2 Drawbacks

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

2.3 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.

2.3.1 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. The enrichment in discrimination reduction with the relabeling method is very satisfying. 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.

2.3.2 Drawbacks

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

2.4 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 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 dataset contains the discriminatory items. The utility measures of indirect discrimination is same as the above preprocessing approach based on the redlining rule dataset RR

2.4.1 Result

The values of DDP and DPD achieves a high degree of indirect discrimination prevention in different cases. 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.

2.4.2 Drawbacks

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

2.5 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 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.

2.5.1 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 we should enforce the following inequality for each redlining ruler: $D,B \rightarrow C$in RR: *elb*

$(\gamma, \delta) < \alpha$ These two data transformation method for used simultaneous direct and indirect discrimination prevention.

2.5.2 Utility Measures

These techniques should be evaluated based on two aspects

- 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

2.5.3 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. PROPOSED SOLUTION

The main negative impacts of data mining is discrimination and privacy. 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 completed a wide overview of the distinctive methodologies for discrimination prevention for data mining. We discussed the issues and limitation of the recent state of the approaches. Based on the same issues, we study an approach that uses transformation method. This approach helps to prevent direct discrimination and indirect discrimination.

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.