International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Discrimination Prevention in Data mining with Privacy Preservation

Anjali P S[1], Renji S[2]

[1]CUSAT, Sarabhai Institute of Science and Technology, Vellanad, Trivandrum, (DIST)-Kerala 695543, India

[2]CUSAT, Sarabhai Institute of Science and Technology, Vellanad, Trivandrum, (DIST)-Kerala 695543, India

Abstract: ***Data mining is an important technology for extracting useful knowledge hidden in large collections of data. There are many positive aspects for data mining, but it also has some disadvantages which include potential discrimination and potential privacy invasion. Discrimination can be defined as treating people unfairingly based on their belonging to a particular group. On the other hand, privacy invasion is the possibility of learning private personal data by unauthorized people. For automated decision making, the classification rules are learned from the training datasets. If the training datasets are biased according to what is discriminatory or sensitive, discriminatory decisions may occur. Discrimination is of two types such as direct discrimination and indirect discrimination. Direct discrimination occurs when decisions are taken based on discriminatory attributes such as religion, race, etc. Indirect discrimination occurs when decisions are made based on attributes that are highly related with the sensitive ones.Anti-discrimination techniques are adopted to avoid or eliminate the discrimination in data mining. While data mining is done, the private data's in the dataset can be learned by malicious people. To handle this problem, privacy preservation algorithms are used. In this paper, the potential discrimination and privacy invasion is tackled and propose new methods for direct and indirect discrimination prevention and privacy preservation. Also, we discuss the methods for cleaning the datasets and propose metrics to evaluate the utility of proposed approach. The experimental evaluation is done which demonstrates that the proposed antidiscrimination techniques are effective while preserving the privacy of the datas simultaneously.***

Keywords: Data mining, antidiscrimination, direct and indirect discrimination, rule protection, rule generalization, privacy preservation, slicing.

1. Introduction

In data mining, discrimination and privacy are the two of the issues discussed in the recent literature. Discrimination is the denying some advantages to members on their belonging to a particular group. Laws are designed by government to eliminate discrimination to some extent. Discrimination occurs based on attributes such as age, gender, religion, etc. Privacy is the possibility of unauthorized persons accessing the private data. Privacy is one of the main issue while using technology for transferring data. Today, a large amount of datas are collected routinely by banks, insurance companies for loans etc. These data are collected for making a decision on whether to approve a loan or deny it. Automated decision making is done using classification rules mining and data miningdiscriminatory items, discriminatory decisions may occur. Discovering such potential biases, and eliminating them from the training data without harming their decision making utility is therefore highly desirable. Discrimination can be either of the two types. Direct discrimination consists of rules or procedures that explicitly mention minority groups based on sensitive attributes related to the group .Indirect discrimination consists of rules or procedures that while not explicitly mentioning discriminatory attributes intentionally or unintentionally could generate discriminatory decisions. Also personal data are collected for these purposes. So maintaining the privacy of these data are very much important. A new technique developed for privacy preservation is the slicing algorithm. It provides better utility than generalization and preserves the attributes relation than bucketization. Slicing algorithm consists of three phases such as attribute partitioning, column generalization and tuple partitioning.

In this paper, we review the issues of discrimination (both direct and indirect) and privacy. The rest of the paper is organized as follows. The section 2 discussed the existing literature review. Section 3 discussed the analysis of the existing approaches. Section 4 presents a system architecture of the new approach. Section 5 presented algorithmfor discrimination prevention and privacy preservation. At the end, results and conclusion is presented in section 6 and 7.

2. Literature Review& Related Work

This section discusses the state of the art approaches dealing with discrimination prevention and privacy preservation in data mining.

[2] Investigated three approaches for removing discrimination from a Naïve Bayes Classifier. The modifying naïve Bayes method, the observed probabilities in a naïve Bayes model is changed in such a way that its predictions become discrimination free. The second method called the two naïve Bayes models involved learning two models and balancing these models afterwards. And the latent variable model introduced a latent variable reflecting the latent true class of an object without discrimination. All these methods performed the classification of the data in such a way that focuses on independent sensitive attribute and does not consider numerical attribute as a sensitive attribute. [3] and [4] proposed an approach which focusses on the concept of classification without discrimination. In [3], they proposed a solution based on massaging the data to remove the discrimination from it with least possible changes. [4]Classification with No Discrimination by Preferential Sampling guarantees hopeful results with both stable and unstable classifiers. The foremost inspiration behind the

Preferential Sampling (PS) is that the data objects which are close to the borderline have more promise to get discriminated and those data will get high inclination while sampling. This arrangement of data objects ensures that if the rank of the element is high, then it is more close to the borderline. PS starts from the original training dataset and iteratively duplicates and removes objects. . [4]Presented the issues of discrimination in a social sense that is against the minorities and disadvantaged groups. It also attempts to handle a dataset of decision records and uses a classification rule for solving problem. [5]The new idea of constructing the decision trees with non-discriminatory constraints is a divergent to the earlier approaches. As they aims in "removing" undesired dependencies from the training data and thus can be considered as pre-processors. Two approaches namely, Massaging and Reweighing are used to clean away the data.[6]Anti-discrimination also plays a foremost role in cyber security, in which the computational intelligence technologies such as data mining may be used for different decision making scenarios. It is the former work that applies antidiscrimination in cyber security. The main concern here is to deal the problem without corrupting the efficacy of data for cyber security applications that rely upon data mining, e.g. intrusion detection systems. [7]It introduces anti-discrimination in the perspective of cyber-security. It analyzes a new discrimination prevention method based on data transformation that can consider numerous discriminatory attributes and their combinations. But the disadvantage was that they does not run on real datasets and also do not consider the background knowledge. [8]The problem of discrimination-aware classification can be identified by constructing a decision tree classifier without discrimination R. Agarwal and R. Srikant [10] discussed the method of associations rule mining for large databases. Two algorithms were designed that help to discover the association between items in a large database of transactions. However, they did not consider the quantities of the items brought in a transaction, Also, as the problem size increased, the performance gap also increased. [13] described the architecture of DCUBE, and demonstrated the issues of discrimination discovery, by making people aware of the legal issues data can hide, and to an approach for discrimination analysis also guided the audience through the processes for discovering direct discrimination, indirect discrimination, respondent argumentation, affirmative actions and favoritism and allowed participants to directly interact by posing specific queries over the DCUBE database.[14]discusses about the various algorithms for privacy preservation.

3. Analysis of the Problem

During the investigation of literature survey, some issues were identified and are summarized using the following points:

- The relationship between discrimination prevention and privacy preservation in data mining is not explored. It remains unseen whether privacy protection can help anti-discrimination or vice-versa.
- The methods focus on the attempt to detect discrimination in the original data only for one discriminatory item and also based on a single measure
- They do not include any measure to evaluate how much discrimination has been removed and how much information loss has been incurred.
- The synergies between rule hiding in privacy preserving data mining and rule hiding for discrimination removal is not found out.
- It focusses either on direct discrimination or indirect discrimination or not on both together.
- The approaches do not shows any measure to evaluate how much discrimination has been removed, and thus do not concentrate on the amount of information loss generated.

So the proposed work in data mining propose preprocessing methods which overcome the above limitations. And introduces new data transformation methods (rule protection and rule generalization (RG)) are based on measures for both direct and indirect discrimination and can deal with several discriminatory items.

4. Proposed Work

Data mining is the valuable technology for extracting knowledge underlying in large storage of data. Discrimination is one of the destructive effects of data mining. The intention of this system was to enlarge a new pre-processing discrimination prevention methodology including different data transformation methods that can stop direct discrimination, indirect discrimination or together at the same time. To accomplish this idea measure discrimination and individuals which have been directly and/or indirectly discriminated in the decision-making processes ought to be identified. Datas are transformed in the proper way to take out all those discriminatory attributes. Even though there exists more than a few methods for each of the above mentioned approaches, discrimination prevention still remains a largely unexplored research avenue. It aim principally on discrimination prevention based on pre-processing. The pre-processing approach seems the most flexible one, since there is no need to modify the benchmark data mining algorithms. It includes not only knowledge publishing, but also data publishing. The system resolves direct and indirect discrimination either independently or together at the same time.

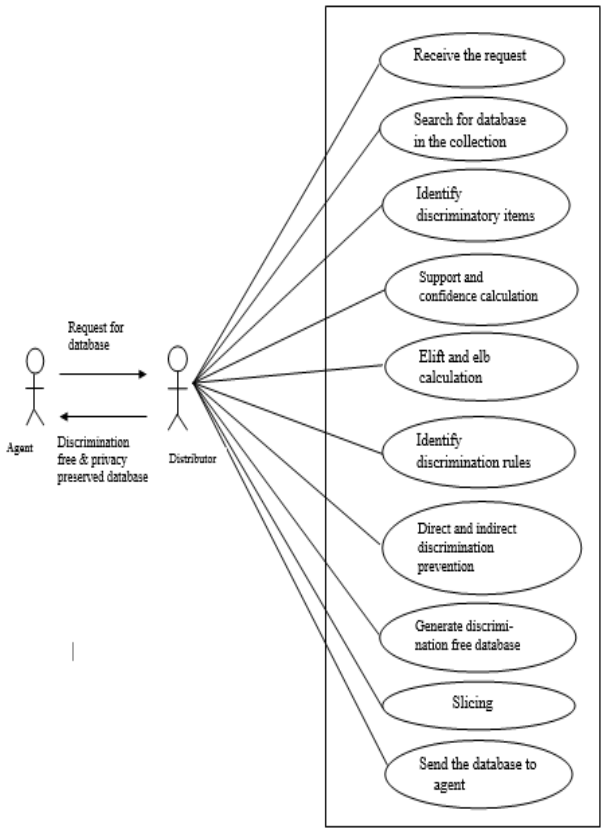

Figure 1: Use case diagram for discrimination prevention and privacy preservation

5. Discrimination Prevention and Privacy Preservation

Firstly the agent has be an authorized service provider or a person. The agent requests for the database as per his need from the distributor, which is discrimination free and privacy preserved. The resultant database is received by the agent in text format. Then he has to retrieve the original database. The distributor collects all the databases as far as possible. When an agent's request for a database is received, the distributor searches for the database in its collection. After finding out the required database, the distributor performs the necessary processes to remove the discrimination and performs privacy preserving based on the level of user. The resultant database is send to the agent in the text format due to security reasons. The process for discrimination prevention consists of two phases such as discrimination measurement and data transformation. On this discrimination free database, the slicing algorithm is used for privacy preservation.

A.*Discrimination measurement and data transformation*

Direct and indirect discrimination discovery includes identifying α-discriminatory rules and red lining rules. Based on the predetermined discriminatory items in database, frequent classification rules are classified in to 2 groups such asPotentially Discriminatory classification rules (PD) and Potentially Non-Discriminatory (PND).If we have a set of classification rules and discriminatory items in database, then a classification rule is said to be PD, when it contains a non-empty discriminatory item set and a non-discriminatory item set.A classification rule is PND, when an item set is nondiscriminatory. A PND rule could lead to discriminatory decisions in combination with some background knowledge.Direct discrimination is measured by identifying α- discriminatory rules among the PD rules using a direct discrimination measure (elift) and a discriminatory threshold (α). The purpose of direct discrimination discovery is to identify α-discriminatory rules. They indicate biased rules that are directly inferred from discriminatory items and are called as α-discriminatory direct rules.Indirect discrimination is measured by identifying redlining rules among the PND rules combined with background knowledge; using an indirect discriminatory measure (elb) and discriminatory threshold (α).The purpose of indirect discrimination discovery is to identify redlining rules. Redlining rules indicate biased rules that are directly inferred from non-discriminatory items because of their correlation with discriminatory ones.

The data transformation transform the original database 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 transformed data. Data transformation for direct discrimination

A suitable data transformation with minimum information loss should be applied in such a way that each α-discriminatory rule either becomes α-protective or an instance of redlining rule. The first is called as direct rule protection (DRP) and second one rule generalization (RG).Direct Rule Protection either changes the discriminatory item set in some records or changes the class item in some records. In rule generalization, the relation between the rules is considered instead of discrimination measures. A PD rule is an instance of PND rule, if the PD rule has the same or higher confidence than the PD rule. In data transformation for indirect discrimination, the data set of decision rules would be free of indirect discrimination if it contained no redlining rules. To achieve this, a suitable data transformation with minimum information loss is applied in such a way that the redlining rules are converted in to non-redlining rules. This procedure is called as Indirect Rule Protection (IRP).Here the measure 'elb, is used in order to turn a redlining rule in to a non-redlining rule.

B. Privacy Preservation Using Slicing

Slicing is an efficient algorithm for computing the sliced table that satisfies l-diversity.The algorithm consists of three phases such as attribute partitioning, column generalization and tuple partitioning. Slicing partitions the dataset both horizontally and vertically. Vertical partitioning contains grouping of attributes in to columns based on correlations among the attributes. Horizontal partitioning contains by grouping tuples in to buckets. Within each bucket, values in each column are randomly sorted to break the linking between different columns. Slicing can handle both high dimensional data and data without a clear separation of quasi-identifiers and sensitive attributes.

6. Experimental Result

The ant-discrimination methodology used was Rule protection and Rule generalization. And for privacy preservation, slicing algorithm was used. The entire process for discrimination prevention and privacy preservation in data mining is as per shown in fig 1.German credit dataset [11] was used for experimental purpose, which is a well-known real-life data set, containing both numerical and categorical attributes.

There are several notable findings in this work. The German credit dataset is a large dataset consisting of 1000 records and 20 attributes (without class attribute) of bank account holders. So processing this large dataset in a single stretch was not possible. So this dataset was converted in to a database. Then database was divided in to few sections and processed them simultaneously. The discriminatory items were identified by both the agent and the process. The sensitive attribute was taken as discriminatory by the process and the other discriminatory items were notified by the requester of database .Based on the both, discrimination prevention process was carried out. Also the relationship between discrimination prevention and data mining was explored and was found that the privacy preservation can help discrimination prevention.

7. Conclusion

Data mining is an increasingly important technology for extracting useful knowledge hidden in large collections of data. The negative perceptions of data mining includes potential privacy invasion and potential discrimination. The main aim of this thesis work is to develop a system that provides discrimination prevention as well as privacy preservation. Discrimination prevention in data mining aims at discovering unfair decisions and behavior and preventing taking similar decisions by authorized people. Privacy preserving in data mining aims at preventing the possibility of learning private personal data by unauthorized people. The direct and indirect discrimination prevention consists of two phases such as discrimination measurement and data transformation. In discrimination measurement, the potentially discriminatory and non-discriminatory rules are identified. Then direct and indirect discrimination is measured using elift and elb functions. For data transformation for direct discrimination rules, direct rule protection and direct rule generalization algorithms are implemented. For data transformation for indirect discrimination, indirect rule generalization algorithm is used. After data transformation, the discrimination – free database is obtained. To this discrimination-free database, privacy preserving algorithm called slicing is performed. The resultant database is discrimination free and privacy preserved.

Our analysis and experiments show that there is a relationship between the discrimination prevention and privacy preservation in data mining. By using this relationship, the performance of the system can be increased.

References

[1] Sara Hajian and Joseph Domingo-Ferrer A Methodology for Direct and Indirect Discrimination Prevention in Data Mining Fellow, IEEE, 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 without Discrimination," Proc. IEEE Second Int'l Conf. Computer, Control and Comm. (IC4 '09), 2009.
[6] F. Kamiran and T. Calders, "Classification with no Discrimination by Preferential Sampling," Proc. 19th Machine Learning Conf. Belgium and The Netherlands, 2010.
[7] F. Kamiran, T. Calders, and M. Pechenizkiy, "Discrimination Aware Decision Tree Learning," Proc. IEEE Int'lConf. Data Mining (ICDM '10), pp. 869-874, 2010.
[8] 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.
[9] 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.
[10] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," Proc. 20th Int'lConf. Very Large Data Bases, pp. 487 -499, 1994.
[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] P.N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Addison-Wesley, 2006.
[13] S. Ruggieri, D. Pedreschi, and F. Turini, "DCUBE: Discriminationin Databases," Proc. ACM Int'l Conf. Management of Data (SIGMOD '10), pp. 1127-1130, 2010
[14] 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
[15] P.N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining.Addison-Wesley, 2006
[16] 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