

A Novel Fast and Accurate Method for Preventing Discrimination in Data Mining

¹P Baskaran and ²Arulanandam

¹Dept of Computer Science, Manonmanium Sundaranar University, Tirunelveli, India

²Dept. of Computer Applications, G.T.M Govt. Arts
and Science College, Gudiyatam, Tamilnadu, India

Abstract: Discrimination is a presuppose privileges where provide to the each separate group for the safety of the data which is stored. Discrimination is two type direct and indirect. In direct discrimination is supported sensitive data. Direct discrimination is supported sensitive data. In indirect discrimination is supported unrestricted information. Existing system standard algorithm is used. Sometimes the data should be lost. In this data model sensitive information should be free. It doesn't successfully handle the indirect discrimination problems which are associated to direct discrimination. In the system using the new techniques to prevent the sensitive information. Discrimination deterrence methods in term of data quality and discrimination detection for both direct and indirect discrimination.

Key words: Discrimination • Direct and Indirect Discrimination • Privileges

INTRODUCTION

Discrimination involves the group's initial reaction that influencing the individual's actual behavior towards the group, restricting members of one group from privileges that are available to another group, leading to the rejection of the individual or entities based on logical decision making. Discrimination based on age, religion, gender, caste, disability, employment, language, race and nationality. There are several decision-making tasks which made them to discrimination, e.g. loan granting, education and health insurances. Given a set of information items on a customer, an automated system decides whether the customer is to be recommended for a credit or a certain type of life insurance. Automating such decisions reduces the workload of the staff of banks and insurance companies, among other organizations. Age discrimination is discrimination that depends on beliefs and values which used to justify discrimination and subordination based on someone's age. Ageism defines that it directed towards old people, or adolescents and children. Disability discrimination is the process of individuals which treats as the standard of usual living that ends in public and private places, education and social work that are built to survive best people, thereby rejecting those with various disadvantages.

Denying someone job opportunities, or disallowing one from applying for particular jobs, is often considered as employment discrimination for such a rejection is not related to the requirements protected characteristics may include age, disability, ethnicity, weight, religion, gender, gender identity, height, nationality, gender orientation and skin color. In the beginning, automating decisions may give a sense of fairness but the decision rule does not learn itself by personal preferences. The classification rules are actually learned by the system model based on historical data. If the original data are inherently biased against a particular community, the learned model may also show the negative impact on it. For example, in a certain loan granting organization, foreign people are rejected for loan for the years. If this biased historical dataset is used as training data to learn classification rules for an automated loan granting system, the learned rules also suffered from biased behavior toward foreign people and also the system may consider that the foreign is a legitimate criterion for loan rejection.

Related Work: Numerous direct and indirect discrimination schemes have been proposed previously. Those schemes either eliminate direct or indirect discrimination. Fast algorithms for mining association rules that defines the issues of discovering association

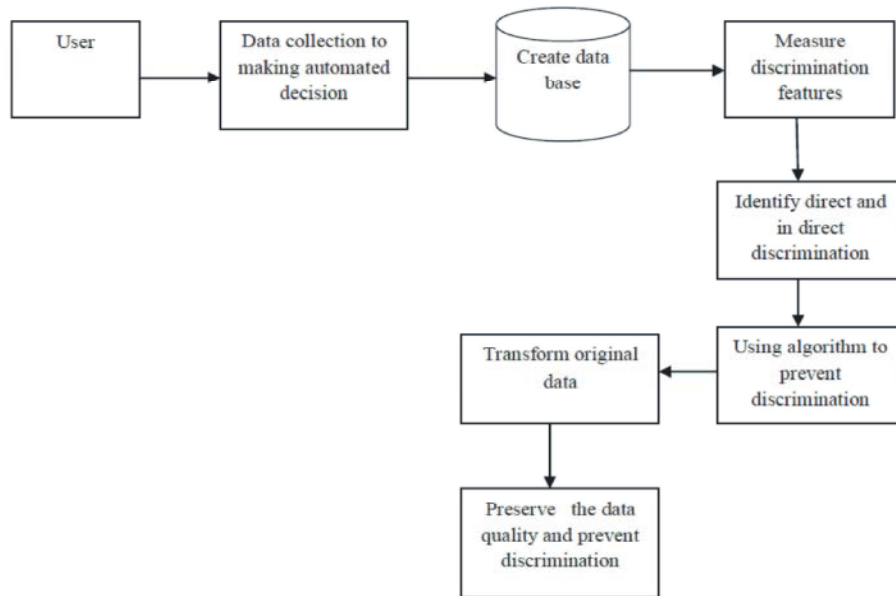


Fig. 1: System Architecture

rules between items in a large database of sales transactions [1]. The proposed algorithms can be combined into a hybrid algorithm named as Apriori Hybrid. Figure 1 describes the process of extracting biased and unbiased decision rules. Data mining with discrimination sensitive and privacy sensitive attributes states that the privacy legitimate input data and the output data are used for selecting preserving protection against such discrimination. In this technique it describes the work in progress of research project based on legal and ethical rules can be combined in data mining algorithms to prevent such activities.

Sara Hajian investigated that how to clean training datasets and outsourced datasets in such a way that legitimate classification rules can still be extracted but discriminating rules based on sensitive attributes cannot [2, 3]. To overcome this issue, anti-discrimination process is introduced. It includes discrimination discovery and prevention techniques. Discrimination discovery is based on mining classification rules and reasoning on them on the basis of quantitative and qualitative measures of discrimination [4-15]. This approach has been extended to encompass statistical significance of the extracted patterns of discrimination and reason about affirmative action. Discrimination prevention represents a set of patterns that do not allow for discriminatory decisions even if the historical data sets are discriminated. It involves three approaches: 1. Pre-processing 2. In-processing 3. Post-processing In this paper, we concentrate on preprocessing techniques.

Classification Based on Discrimination Prevention Using Data Transformation Techniques:

Classification is the task of generalizing known structures applies to new data. Classification is supervised learning. For example, classes are used to represent that a customer defaults on a loan decisions like ‘Yes’ or ‘No’. It is important that each record in the dataset used to represent the classifier already have a value for the attribute used to describe classes. Because each record has the attribute value used to define the classes. Classification is a machine learning technique used to predict group membership for data instances. It assigns items in a collection to target categories. The aim of classification is to accurately determine target class for each and every case in data. Direct discrimination restricts a particular community based on sensitive reasons. Indirect discrimination restricts certain number of peoples based on non sensitive one. 1. DISCRIMINATION MEASUREMENT The purpose of Discrimination measurement is to identify discriminatory rules and redlining rules using Potentially Discriminatory (PD) and potentially non-discriminatory (PND) rules. Direct discrimination is measured by identifying α - discriminatory rules among the PD rules using a direct discrimination measure (elift) and threshold (α) [16-22].

Classification with No Discrimination by Preferential Sampling:

We can remove the sensitive data instead of relabeling it. The new solution to the CND problem by introducing a sampling scheme for making the

discrimination free instead of relabeling the data set. The algorithm is used in this paper is classification algorithm. The goal of classification is to accurately predict the target class for each care in the data. Predicts categorical labels and classify the data based on the training set and the values in a classifying attribute and uses it in classifying new data. The techniques used in this paper is Pre-processing, Preferential sampling, Over sampling, Uniform sampling. In preprocessing there are a lot tangential and excess data present or noisy and so knowledge uncovering during the aiming stage are a lot of elaborated. Data preparation and filtering steps can considerable amount of processing period. Data pre-processing includes cleaning, normalization, transformation and characteristic extraction selection. In Preferential sampling arises when the process that determines the data location and the process being modeled are stochastically dependent. In the over sampling. In the over sample are the action of sample importantly higher than the doubly the twice the band width or peak relative frequency of the signal comprising sampled. Over sampling sample aids avoid aliasing, answer and brings down noise. The equation is used $f_s = 2b$ Where f_s are the sample relative frequency and b are the bandwidth or maximum relative frequency of signal. Then rate is then $2b$. the Uniform sampling defined as each Data objects probability is uniform. In this paper disadvantage is Discrimination were removed in ethical and legal region.

Three Naive Bayes Approaches for Discrimination Free Classification: 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 Naviesbayes model, Latent variable model and Modified naivesbayes. The naivesbayes model is Abayes classifier is a simple possibility classifier based on applying bayes theorem with strong statistical independence assumption. Depending on precise nature of the probability model, naviebayes 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 naviebayes is Modify the probability distribution $p(s/c)$ of the sensitive attribute values given the class values.

Fast Algorithm for Mining Association Rules: Fast algorithm is an efficient algorithm used to avoid the discrimination in data mining. In this paper algorithm apriori, aprioritid, AIS algorithm, apriorihybrid algorithm. The apriori algorithmic rule are the adult detail sets by the early authorize comprised reached aim the fresh candidate detail. Pruning comprised represented applying the information that some subdivision of itemset. The difference for determining the support the database is not used after the first pass. In the AIS algorithm. In the AIS algorithm involves two concepts are extension of an item set, determining what should be in the candidate item set. The apriori hybrid algorithm is Uses apriori in the early passes and later shifts to aprioritid. In this paper disadvantages is An extra cost is sustained when shifting from apriori to aprioritid [10,11].

Discrimination Prevention in Data Mining since Intrusion and Crime Detection: In this paper techniques is used the anti discrimination techniques. Antidiscrimination law to the natural law about the correct by people to comprise addressed. In the governmental involvement people essential comprise addressed on active equal base in some cause by sex activity, age, race, nationality. The approaches are used preprocessing, post processing. The preprocessing is data preprocessing is the important process in the data mining. In there are more than tangential and excess information represent or noisy and uncertain information and then knowledge discovery on the aiming point are heavier The analyzed information that are not represented carefully screened out as much troubles may develop misleading answers. The post processing are the action by categorization through with large amounts of information and cleaning our applicable data. Data mining in acknowledgment to initiative resource preparation are the statistical and ordered analysis from large sets by transaction information mining in relation to enterprise resource planning is the statistical and logical analysis of large sets of transaction data. the algorithm used in this paper is not efficient this is main drawback of this paper.

Visual Data Mining for Higher-level Patterns: Discrimination-Aware Data Mining and Beyond: In this paper, we propose a visualization approach that can on the one hand be applied to any (classification or association) rules, but that is appropriate to bringing out characteristic of mined patterns that are especially important in discrimination-aware and privacy aware data mining. We define new interestingness proceeding for

items and rules and show various ways in which these can help in highlighting information in communicating settings. We conclude by arguing how this approach can lead to a new generation of feedback and awareness tools. The need to inspect mining results carefully for such meta-level relationships between features and outcomes becomes even stronger when specific data, rules and other patterns become the object of scrutiny: The flipside of data mining is that it may make relationships visible that various stakeholders do not wish to become explicit and that the patterns it finds may suggest actions that various stakeholders do not wish to be taken. Such concerns may lead to a new approach to keep and/or treat these data as private.

Proposed System: Figure 2: Process of direct discrimination and indirect discrimination The extended lift of the rule can be calculated as Figure 2 defines the process of direct discrimination and indirect discrimination that generate negative impacts. The indirect discrimination is measured by identifying redlining rules among the PND rules that correlated with background knowledge based on an indirect discriminatory measure (ϵ) and a discriminatory threshold (α). In fact, redlining rules indicate biased rules that are indirectly inferred from nondiscriminatory items because of their correlation with discriminatory ones.

Data Transformation: Transform the original data DB in such a way to remove direct and indirect discriminatory biases, with minimum impact on the datasets. So there is no other negative impact can be discovered from transformed datasets [16]. The data transformation method should increase or decrease the confidence of the rules to the target values with minimum impact on data quality, maximize the disclosure prevention measures and minimize the information loss measures. Data transformation includes rule protection and rule generalization methods for both f-direct and indirect discrimination.

Rule Protection Algorithm for Direct and Indirect Discrimination: In direct discrimination, rule protection algorithm is used to convert each α -discriminatory rule into a α' -protective rule based on the direct discriminatory measure. There are two methods that could be applied for direct rule protection.

Method 1 modifies the discriminatory item set in some records. 2. Method 2 modifies the class item in some records from grant credit to deny credit in the records with male gender. Indirect rule protection is used to turn a

redlining rule into a non-redlining rule, based on the indirect discriminatory measure. Rules that are associated with some background knowledge indirectly called as redlining rules.

CONCLUSION

Along on the privacy, discrimination are an indistinguishable authoritative effect while believing the eligible and right looks from data mining. It is more than noticeable that most people do not want to be discriminated since by their sex, religion, nationality, age and soon, especially when those attributes are applied as constructing conclusions around it alike applying it an job, loan, policy, etc. The aim of these paper was to develop a new pre processing discrimination prevention methodology including different data transformation methods that can prevent direct discrimination indirect discrimination or both of them at them at the same time. To attain this objective, the first step is measure discrimination and identify categories and groups of individuals that have been directly and indirect discriminated in decision making process. The second step is to transform data in proper way to remove all those discrimination biases finally, discrimination free data models can be produced from the transformed data set without seriously damaging data quality

REFERENCES

1. Kamiran F. and T. Calders, 2010. Classification with no Discrimination by Preferential Sampling, Proc. 19th Machine Learning Conf. Belgium and The Netherlands.
2. Calders T. and S. Verwer, 2010. Three Naive Bayes Approaches for Discrimination-Free Classification, Data Mining and Knowledge Discovery, 21(2): 277-292.
3. Hajian S., J. Domingo-Ferrer and A. Martí'nez-Balleste, 2011. Discrimination Prevention in Data Mining for Intrusion and Crime Detection, Proc. IEEE Symp. Computational Intelligence in Cyber Security (CICS '11), 47-54.
4. Pedreschi D., S. Ruggeri and F. Turini, 2008. Discrimination Aware Data Mining, Proc. 14th ACM Int'l Conf. Knowledge Discovery and Data Mining (KDD '08), 560-568.
5. Agrawal R. and R. Srikant, 1994. Fast Algorithms for Mining Association Rules in Large Databases, Proc. 20th Int'l Conf. Very Large Data Bases, 487-499.

6. Calders T. and S. Verwer, 2010. Three Naive Bayes Approaches for Discrimination-Free Classification, *Data Mining and Knowledge Discovery*, 21(2): 277-292.
7. Hajian S., J. Domingo-Ferrer, 2011. Rule Protection for In-direct Discrimination Prevention in Data Mining, *Proc. Eighth Int'l Conf. Modeling Decisions for Artificial In-telligence (MDAI '11)*, 211-222.
8. Kamiran F. and T. Calders, 2009. Classification without Discrimination, *Proc. IEEE Second Int'l Conf. Computer, Control and Comm. (IC4 '09)*.
9. Kamiran F. and T. Calders, 2010. Classification with no Discrimination by Preferential Sampling, *Proc. 19th Machine Learning Conf. Belgium and The Netherlands*.
10. Kamiran, F. and T. Calders, 2010. Discrimination Aware Decision Tree Learning, *Proc. IEEE Int'l Conf. Data Mining (ICDM '10)*, 869-874.
11. Mason, J., 2002. Filtering spam with SpamAssassin, in *Proc. HEANetAnnu. Conf.*
12. Sahami, M., S. Dumais, D. Heckerman and E. Horvitz, 1998. A Bayesian approach to filtering June e-mail. In *learning for text categorization*, in *Proc. Workshop*.
13. Meyer T.A. and B. Whateley,, 2004. Spam Bayes: Effective open-source, Bayesian based, e-mail classification system, in *Proc. CEAS*.
14. MAPS, 1996. RBL-RealtimeBlackhole List [On-line]. Available and Phishing Attacks, *Cryptology ePrint Archive*, Rep. 2004/155 [On-line]. Available:<http://eprint.iacr.org/2004/155>
15. The Spamhaus Project. The Spamhaus Block List [Online]. Available <http://www.spamhaus.org/sbl>
16. Herzberg and A. Gbara, 2004. Trustbar: Protecting (Even Naive) Web Users From Spoofing.
17. Net craft. Anti Phishing Toolbar, <http://www.mail-abuse.com/services/mds-rbl.html>.
18. Herley C. and D. Florencio, 2006. How to login from an Internet cafe without worrying about key loggers, in *Proc. SOUPS*.
19. Wiedenbeck, S., J. Waters, L. Sobrado and J. Birget, "Design and evaluation of a shoulder-surfing resistant graphical password scheme, in *Proc. Working Conf. Adv. Vis. Interfaces*.
20. Kohavi R. and B. Becker, 1996. UCI Repository of Machine Learning Data base, <http://archive.ics.uci.edu/ml/datasets/Adult>.
21. Sara Hajian and Josep Domingo-Ferrer, 2013. A Methodology for direct and indirect discrimination in data mining. *Knowledge and Data Engineering*, 25(7): 1445-1459.
22. Agrawal, R. and R. Srikant, 1994. Fast Algorithms for Mining Association Rules in Large Databases, *Proc. 20th Int'l Conf. Very Large Data Bases*, pp: 487-499.