"An NLP Method for Discrimination Prevention Using both Direct and Indirect Method in Data Mining"

¹ Sampada U. Nagmote, ² Prof. P. P. Deshmukh

P.R.Patil College of Engg & Technology, Amravati Sampada.nagmote@rediffmail.com P.R.Patil College of Engg & Technology, Amravati pranjalideshmukh_it@rediffmail.com

Abstract

Today, Data mining is an increasingly important technology. It is a process of extracting useful knowledge from large collections of data. There are some negative view about data mining, among which potential privacy and potential discrimination. Discrimination means is the unequal or unfairly treating people on the basis of their specific belonging group. If the data sets are divided on the basis of sensitive attributes like gender, race, religion, etc., discriminatory decisions may ensue. For this reason, antidiscrimination laws for discrimination prevention have been introduced for data mining. Discrimination can be either direct or indirect. Direct discrimination occurs when decisions are made based on some sensitive attributes. It consists of rules or procedures that explicitly mention minority or disadvantaged groups based on sensitive discriminatory attributes related to group membership. Indirect discrimination occurs when decisions are made based on nonsensitive attributes which are strongly related with biased sensitive ones. It consists of rules or procedures that, which is not explicitly mentioning discriminatory attributes, intentionally or unintentionally, could generate decisions about discrimination.

1. Introduction

Discrimination is a process of unfairly treating people on the basis of their belonging to some a specific group. For instance, individuals may be discriminated because of their race, gender, etc. [5] or it is the treatment to an individual based on their membership in particular category or group. There are various laws which are used to prevent discrimination on basic of various attributes such as race, religion, nationality, disability and age.

There are two types of discrimination i.e. Direct Discrimination and Indirect Discrimination. Direct Discrimination is direct discrimination which consists of procedure or some decided rule that mention minority or disadvantaged group based on sensitive attributes to they are related to membership of group. Indirect Discrimination is discrimination which consists of rules and procedures that are not mentioning attributes which causes discrimination and hence it generates discriminatory decision intentionally or unintentionally.

2. What is Discrimination?

Discrimination should be defined as one person, or a group of persons, being treated less favourably than another on the grounds of racial or ethnic origin, religion or belief, disability, age or sexual orientation (direct discrimination), or where an apparently neutral provision is liable to disadvantage a group of persons on the same grounds of discrimination, unless objectively justified (indirect discrimination).

In other words, discrimination means treating people differently, negatively or adversely without a good reason. As used in human rights laws, discrimination means making a distinction between certain individuals or groups based on a prohibited ground. The idea behind it is that people should not be placed at a disadvantage simply because of their racial and ethnic origin, religion or belief, disability, age or sexual orientation. That is called discrimination and is against the law.

3. Related Work:

Despite the wide deployment and developed of information systems based on data mining technology in decision making, the issue of antidiscrimination in data mining did not receive much attention up to 2008. Some proposals deals with discovery of discrimination and some deals with the prevention of discrimination are discovered. The discovery of discriminatory decisions was first proposed by Pedreschi et al. The approach is based on mining classification rules (the inductive part) a reasoning on them (the deductive part) on the basis of quantitative measures of discrimination that formalize legal definitions of discrimination. For instance, the US Equal Pay Act states that: "a selection rate for any race, sex, or ethnic group which is less than four fifth of the rate for the group with the highest rate will generally be regarded as evidence of adverse impact." This approach has been extended to encompass statistical significance of the extracted patterns of discrimination in and to reason about affirmative action and favoritism. Moreover it has been implemented as an Oracle-based tool in. Current discrimination discovery methods consider each rule individually for measuring discrimination without considering other rules or the relation between them. Discrimination prevention is used for preventing discrimination without affecting the data quality. Discrimination prevention can be done by using three approaches,

A) **Pre-processing:-** In this approach we used hierarchical based generalization, and transform the data in such a way that biased data are removed so that no unfair decision rule can be mined from the transformed data.

B) In-Processing: - In this approach the actual data mining algorithms is changes in such a way that the resulting models do not contain unfair decision rules. For example, an alternative approach to cleaning the discrimination from the original data set is proposed in whereby the nondiscriminatory constraint is embedded into a decision tree learner by changing its splitting criterion and pruning strategy through a novel leaf relabeling approach.

C) Post-processing:- In this approach the resulting data mining models is modify, instead of cleaning the original data set or changing the data mining algorithms. For example, in, a confidence-altering approach is proposed for classification rules inferred by CPAR algorithm.

4. Contribution and plan of this paper:

Discrimination prevention methods based on preprocessing published so far present some limitations, which we next highlight. They attempt to detect discrimination in the original data only for one discriminatory item and based on a single measure. This approach cannot guarantee that the transformed data set is really discrimination free, because it is known that discriminatory behaviors can often be hidden behind several discriminatory items, and even behind combination of them. They only consider direct discrimination. They do not include any measure to evaluate how much discrimination has been removed and how much information loss has been incurred. In this paper, we propose preprocessing methods which overcome the above limitations.

In this paper we are going to use Natural Language Processing Approach for direct and indirect discrimination prevention. It consists of POS tagging and chunking methods. POS tagging is useful for identifying verbs, nouns, adjectives in a given line. On the basis of that we can identify the action words which may cause direct or indirect discrimination.



Figure1

In the figure 1, we have to select the one file in the form of text file, pdf file & csv form data. After the selection of file then the operation such as split, tokenize, Pos tag, chunk operation is performed.



Figure 2

In the figure 2, splitting operation is performed in this splitting operation each and every sentence from the file is separated.





Figure 6

In the figure 6, the actual result got in the form of discriminatory words found, indiscriminatory words found, % of discriminatory words,% of indiscriminatory words, time required for all this processes and final document.



Figure 7

In the figure 7, Admin can add words from his point view causes discrimination and this words stores in the dataset.

5. Conclusion

As along with privacy, discrimination is very important issue considering legal and ethical aspects of data mining. As person would like to be get discriminated on the basis of their gender, age, locality, religion and so on, especially when those attributes have a big impact on decision making about them like loan, job, and Insurance etc.

The purpose of this paper is to identify direct, indirect or both type of discrimination and prevent them without losing the data quality. For this purpose we are going to use NLP and its tasks.

6. References:

[1] Sara Hajion & josep Domingo-Ferrer, Fellow "A Methodology for Direct and Indirect Discrimination Prevention in Data Mining," Knowledge and Data Engineering

[2]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.

[3] European Commission, "EU Directive 2004/113/EC on Antidiscrimination," http://eur-

Figure 3

In the figure 3, the tokenize operation is performed. In this form each and every word from the file is separated.



Figure 4

In the figure 4, the operation called POS tag is performed. In this Pos tag method separate the each noun, pronoun, adjective, adverb phrases.



Figure 5

In the figure 5, the chunking operation is performed. In this action words are removed and the remaining words are compare with dataset .During comparison the common words are removed. lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:373:003 7:0043:EN:PDF, 2004.

[4] European Commission, "EU Directive 2006/54/EC on Antidiscrimination," http://eur-lex.europa.eu/LexUriServ/

LexUriServ.do?uri=OJ:L:2006:204:0023:0036:en:PDF, 2006.

[5] 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.

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

[7] F. Kamiran and T. Calders, "Classification without Discrimination," Proc. IEEE Second Int'l Conf. Computer, Control and Comm.(IC4 '09), 2009.

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

[9] 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.

[10] R. Kohavi and B. Becker, "UCI Repository of MachineLearningDatabases,"http://archive.ics.uci.edu/ml/datasets/Adult, 1996.

[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] 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.

[13] 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.

[14] D. Pedreschi, S. Ruggieri, and F. Turini, "Integrating Induction and Deduction for Finding Evidence of Discrimination," Proc. 12th ACM Int'l Conf. Artificial Intelligence and Law (ICAIL '09), pp. 157-166, 2009.

[15] 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.

[16] 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.

[17] P.N. Tan, M. Steinbach, and V. Kumar, Introduction to DataMining. Addision-Wesley 2006.