ANALYSIS

4(12), June 1, 2013

ISSN 2278–5469 EISSN 2278–5450

Discovery

An analysis of discrimination discovery and prevention for recruiting employees from social networks and e-jobs

Arulanandam K[1], Baskaran P[2☼]

1.Reaearch Supervisor, Professor, Dept. of Computer Applications, Manonmanium Sundaranar University, Tirunelveli, India
2.Research scholar, Dept of Computer Science, Manonmanium Sundaranar University, Tirunelveli, India

☼Corresponding author: Research scholar, Manonmanium Sundaranar University., Tirunelveli, India, E-mail: Mail: bas_vlr@yahoo.co.in

Publication History
Received: 12 April 2013
Accepted: 19 May 2013
Published: 1 June 2013

Citation
Arulanandam K, Baskaran P. An analysis of discrimination discovery and prevention for recruiting employees from social networks and e-jobs. *Discovery*, 2013, 4(12), 55-59

ABSTRACT

Data mining is a concept of various techniques and collection of algorithms for extracting knowledge from large collections of various data sources. But however, a bad and harmful social encasements about data mining, among which potential privacy assault and potential discrimination. The concluding consists of wrongly and unjustified treating people on the basis of their belonging to a specific group. And also the act of making generalized distinctions among groups of people or things without inquiry into the specific characteristics of individuals or within the group and also includes cyber frauds. Social networking is the process of finding friends and of managing friendships through the internet. People who wish to meet others online set up and about their most convincing and eye-catching presentations through their profile pages. They join groups and communicate with others by commenting on topics or by introducing topics that hope to encourage discussion. Mining algorithms are training from datasets which may be prejudiced in what regards gender, race, religion or other attributes.., discriminatory decisions may precede. For this reason, anti-discrimination techniques including discrimination discovery and prevention for various members attitude in social network have been introduced in data mining. We deal with discrimination prevention in data mining and propose new techniques relevant for discrimination prevention individually or group at the same time I social network. In this analysis, we discuss how to clean training datasets and outsourced datasets in such a way that lawful classification rules can still be extracted but discriminating rules based on sensitive attributes cannot. The experimental evaluations demonstrate that the proposed techniques are effective at removing direct and/or indirect discrimination biases in the original dataset while preserving data quality of dataset.

Keywords: anti-discrimination, knowledge discovery, cyberfaurds, datasets, prevention.

1. INTRODUCTION

1.1. Deprivation of social networks

Everything that is an advantage about social networking can also be a disadvantage in that you lose your privacy - after all, you have volunteered personal information that is now online. Every site allows you to set privacy settings. These can be changed from their default settings to limit what other people can see and read about you. For example, you could set your pages to be only viewed by friends. Other parts could be made public. Other parts could be set to family only. Some disadvantages are

- You lose some privacy compared to not being on a social network

- You may later regret posting pictures or comments that you thought funny at the time
- Online bullying can be a problem if someone posts unkind or untrue things about you
- Some people may use a fake profile - just because they say they are 15 years old does not mean that is true. Be careful when you choose to be friends with someone you have never met in real life
- They can be a real distraction and time waster, some people spend many hours on social networking rather than be working or studying. For example, constantly checking their twitter feeds.
- Take everything you see with a pinch of salt - people do like to boast and overstate just like they do in real life.

2. A CASE STUDY

Social networking websites are fast becoming a staple of corporate recruiting. Depending on which studies you read, anywhere from 39 to 65 percent of companies use social networking websites to identify and screen potential candidates for open positions. Sites like LinkedIn, Facebook, Twitter and Ning have made it easier and cheaper for recruiters and hiring managers to access a vast and receptive talent pool, Some peoples from corporate as a HR consultant who specializes in social media notes that there are 600 million active users on Facebook alone who spend between six and 12 hours each month on the site. In addition, these sites can offer recruiters a view into candidate's personalities and work styles that they may never otherwise get from a resume, cover letter or job interview. The Web is used to find candidates for retail jobs while working. And also to see some dating websites, local, city chat rooms and community forums to source candidates. This web-based sourcing strategy worked well for Target, but later, we see most of the candidates came from Facebook and MySpace; job seekers in particular had a higher retention rate as opposed to hiring someone from a job fair or newspaper. But the benefits that social networking websites offer to recruiters and hiring managers in terms of the information they provide about their members also poses a hugee legal risk. Because of the way people meld the personal and the professional on these sites, hiring managers who use them risk factoring inappropriate information about a candidate that they learn through one of these sites into a hiring decision. A hiring manager checking out a candidate's Twitter feed might find out that the candidate has a health condition. The hiring manager, concerned that the candidate will miss a lot of work or cause the company's health insurance premiums to rise, may pass on the candidate, which is a form of illegal discrimination, according to the Americans with Disabilities Act and Title VII of the Civil Rights Act of 1964.

2.1. The Data Protection Act and Data Discrimination

The Data Protection Act controls how your personal information is used by organizations, businesses or the government. Everyone who collects data has to follow strict rules called 'data protection principles'. They must make sure the information is:

- used fairly and lawfully
- used for limited, specifically stated purposes
- used in a way that is adequate, relevant and not excessive
- accurate
- kept for no longer than is absolutely necessary
- kept safe and secure
- not transferred outside the UK without adequate protection

There is stronger legal protection for more sensitive information, such as:

- ethnic background
- political opinions
- religious beliefs
- health
- sexual health
- criminal records

Data Discrimination is a comparison of the general features of target class data objects with the general features of objects from one or a set of contrasting classes. For example, a data mining system should be able to compare two groups of colleges such as the colleges getting a result of 80% distinction and some colleges rarely reaching that mark. Data discrimination is the selective filtering of information by a service provider. This has been a new issue in the recent debate over network neutrality. Accordingly one should consider net neutrality in terms of a dichotomy between types of discrimination that make economic sense and will not harm consumers and those that constitute unfair trade practices and other types of anticompetitive practices. Non-discrimination mandates that one class of customers may not be favored over another so the network that is built is the same for everyone, and everyone can access it.

3. DISCOVERING DISCRIMINATION

Discrimination discovery is about finding out discriminatory decisions hidden in a dataset of historical decision records. The basic problem in the analysis of discrimination, given a dataset of historical decision records, is to quantify the degree of discrimination suffered by a given group (e.g. an ethnic group) in a given context with respect to the classification decision (e.g. intruder yes or no). Figure shows the process of discrimination discovery, based on approaches and measures described in this section.

3.1. Basic Definitions

• An item is an attribute along with its value,
e.g.{Experiance=Fresher}.

• Association/classification rule mining attempts, Given a set of transactions, to predict the occurrence of an item based on the occurrences of other items in the transaction.

• An itemset is a collection of one or more items, e.g.
{Experience=5, Gender=Male}.

• A classification rule is an expression I →CI, where I is an item set, containing no
Class items and CI is a class item,
e.g.{Experience=5, Gender=Male} → Intruder=YES. I is called the premise (or the body) of the rule.

• The support of an itemset, supp(I), is the fraction of records that contain the itemset I. We say that a rule I → CI is completely supported by a record if both I and CI appear in the record.

• The confidence of a classification rule, conf(I →CI), measures how often the class item C appears in records that contain I.
Hence, if supp(I) > 0 conf(I → CI) = supp(I,CI)/supp(I)
Support and confidence range over [0, 1]. In addition, the notation also extends to negated item sets, i.e. ¬I.

• A frequent classification rule is a classification rule with a support or confidence greater than a specified lower bound. Let DB be a database of original data records and FRs be the database of frequent classification rules.

3.2. Potentially Discrimination and Non-Discrimination Classification Rules

With the assumption that discriminatory items in DB are predetermined
(e.g. Experience=5, Gender=), rules fall into one of the following two classes with respect to discriminatory and non-discriminatory items in DB.

1) A classification rule I → CI is potentially discriminatory (PD) when I = A,B with A a non-empty discriminatory itemset and B a non-discriminatory itemset. For example {Experience=5, Gender=Male} → Intruder=Yes.

2) A classification rule I → CI is potentially non-discriminatory (PND) when I is a non-discriminatory itemset.

For example {Experience=5, Gender=Male} → Intruder=YES. The word "potentially" means that a PD rule could probably lead to discriminatory decisions, so some measures are needed to quantify the discrimination potential. Also, a PND rule could lead to discriminatory decisions if combined with some background knowledge, e.g. if in the above example one knows that zip 43700 is mostly inhabited by black people (indirect discrimination).

3.3. Discrimination Measures

Pedreschi et al. (2008), and Verykios et al. (2004) translated the qualitative statements in existing laws, regulations and legal cases into quantitative formal counterparts over classification rules and they introduced a family of measures of the degree of discrimination of a PD rule. In our contribution we use their *extended lift* measure (*elif t*), which is recalled next.

Definition 1: Let $X,Y \rightarrow CI$ be a classification rule
with $conf(Y \rightarrow CI) > 0$. The extended lift of the rule is
$elif\ t(X,Y \rightarrow CI) =$
$conf(X,Y \rightarrow CI)$
$conf(Y \rightarrow CI)$

The idea here is to evaluate the discrimination of a rule by the gain of confidence due to the presence of the discriminatory items (*i.e. X*) in the premise of the rule.

Indeed, *elif t* is defined as the ratio of the confidence of the two rules: *with* and *without* the discriminatory items. Whether the rule is to be considered discriminatory can be assessed by thresholding2 *elif t* as follows.

Definition 2: Let $\alpha \in R$ be a fixed threshold. A PD
classification rule $c = A,B \rightarrow C$ is α-protective w.r.t.
elif t if $elif\ t(c) < \alpha$. Otherwise, c is α-discriminatory.
Consider rule
$c = \{$Experience=5, Gender=Male$\} \rightarrow$ Intruder=YES
If $\alpha = 1.4$ and $elif\ t(c) = 1.46$.

In terms of indirect discrimination, the combination of PND rules with background knowledge probably could generate α-discriminatory rules. If a PND rule c with respect to background knowledge generates an α-discriminatory rule, c is an α-discriminatory PND rule and, if not, c is an α-protective PND rule. However, in our proposal we concentrate on direct discrimination and thus consider only α-discriminatory rules and assume that all the PND rules in *PRs* are α-protective PND. let *MRs* be the database of α-discriminatory rules extracted from database *DB*.

- Note that α is a fixed threshold stating an acceptable level of discrimination according to laws and regulations.

3.4. A Proposal for Discrimination Prevention

In this section we present a new discrimination prevention method which follows the preprocessing approach mentioned above. The method transforms the source data by removing discriminatory biases so that no unfair decision rule can be mined from the transformed data. The proposed solution is based on the fact that the dataset of decision rules would be free of discriminatory accusation if for each α-discriminatory rule $r\_$ there would be at least one PND rule r leading to the same classification result as $r\_$. Our method makes use of the p-instance concept, formalized in the following way.

Definition 3: Let $p \in [0, 1]$. A classification rule $r\_$:

$X,Y \rightarrow CI$ is a p-instance of $r : D,Y \rightarrow CI$ if
1) $conf(r) \geq p \cdot conf(r\_)$ and
2) $conf(r\_\_ : X,Y \rightarrow D) \geq p$.
If each $r\_$ in *MRs* was a p-instance (where p is 1 or a value near 1) of a PND rule r in *PRs*, the dataset of decision rules would be free of discriminatory accusation.

Consider rules r and $r\_$ extracted from the dataset in Table I:

$r_ = \{$Experience=5, Gender=Male$\} \rightarrow$ Intruder=YES
$r = \{$Experience=5, Gender=Male$\} \rightarrow$ Intruder=YES
With $p = 0.8$, rule $r\_$ is 0.8-instance of rule r if:

1) $conf(r) \geq 0.8 \cdot conf(r\_)$
2) $conf(r\_\_) \geq 0.8$

where rule $r\_\_$ is: $r__ = \{$Experience=5, Gender=Male$\} \rightarrow$ PortScan=Yes
Although $r\_$ is α-discriminatory based on the *elif t* measure, the existence of a PND rule r that leads to the same result as rule $r\_$ and satisfies both Conditions (1) and (2) of Definition 3 demonstrates that the subscriber is classified as intruder not because of race but because of using port scanning. Hence, rule $r\_$ is free of discriminatory accusation, because the IDS could argue that $r\_$ is an instance of a more general non-discriminatory rule r. Clearly, r is legitimate, because port scanning can be considered an unbiased indicator of a suspect intruder. Our solution for discrimination prevention is based on the above idea. We transform data by removing all evidence of discrimination appeared in form of α- discriminatory rules. These α-discriminatory rules aredivided into two groups: α-discriminatory rules such that there is at least one PND rule leading to same result and α-discriminatory rules such that there is no such PND rule. For the first group a suitable data transformation with minimum information loss should be applied for ensuring Conditions (1) or (2) of Definition 3 in case they are not satisfied. For the second group, also a suitable data transformation with minimum information loss should be applied in such a way that those α-discriminatory rules are converted to α-protective rules based on the definition of the discriminatory measure

3.5. The detailed process of our solution is described by means of the following phases

• *Phase 1*. Use Pedreschi's measures on each rule to discover the patterns of discrimination emerged from the available data.
• *Phase 2*. Based on Definition 3, find the relationship between α-discriminatory rules and PND
rules extracted in the first phase and determine the transformation requirement for each rule.
• *Phase 3*. Transform the original data to provide the transformation requirement for each respective α-discriminatory rule without seriously affecting the data or other rules.
• *Phase 4*. Evaluate the transformed dataset with the discrimination prevention and information loss measures of Section V-B below, to check whether they are free of discrimination and useful enough. The first phase consists of the following steps. In the first step, frequent classification rules are extracted from *DB* by well-known frequent rule extraction algorithms such as Apriori. In the second step, with respect to the predetermined discriminatory items in the dataset, the extracted rules are divided into two categories: PD and PND rules. In the third step, for each PD rule, the *elif t* measure is computed to determine the collection of α-discriminatory rules saved in *MRs*. The second phase is summarized next. In the first step of this phase, for each α-discriminatory rule in *MRs* of type $r\_ : X,Y \rightarrow CI$, a collection of PND rules in *PRs* of type $r : D,Y \rightarrow CI$ is found. Call *Dpn* the set of these PND rules. Then the conditions of Definition 3, for a value of p at least 0.8, are checked for each rule in *Dpn*. Three cases arise depending on whether Conditions

(1) and (2) hold:
1) There is at least one rule in *Dpn* such that both Conditions (1) and (2) of Definition 3 hold;
2) There is no rule in *Dpn* satisfying both Conditions (1) and (2) of Definition 3, but there is at least one rule satisfying one of those two conditions;
3) No rule in *Dpn* satisfies any of Conditions (1) or (2).

In the first case, it is obvious that currently there is at least one rule r in *Dpn* such that $r\_$ is p-instance of r for $p \geq 0.8$. In this case no transformation is required. In the second case, the PND rule *rb* in *Dpn* should be selected which requires the minimum data transformation to fulfill both Conditions (1) and (2). A smaller difference between the values of the two sides of Conditions (1) or (2) for each r in *Dpn* indicates a smaller required data transformation. In this case, Conditions (1) and (2) in *rb* determine the transformation requirement of $r\_$. The third case happens when there is no r rule in *Dpn* satisfying any of Conditions (1) or (2). In this case, the transformation requirement of $r\_$ determines that this α-discriminatory rule should be converted to an α-protective rule based on the definition of the respective discriminatory measure (*i.e. elif t*). The output of the second phase is a database *T Rs* with all $r\_ \in MRs$, their respective transformed rule *rb* and their respective transformation requirements (see below). The following list shows the first, second and third transformation requirements that can be generated for each $r\_ \in MRs$ according to the above cases:

1) $conf(r\_ : X,Y \rightarrow CI) \leq conf(r : D,Y \rightarrow CI)/p$
2) $conf(r'' : X,Y \rightarrow D) \geq p$
3) If $f() = elif\ t$, $conf(r\_ : X,Y \rightarrow CI) < \alpha \cdot conf(Y \rightarrow CI)$

For the α-discriminatory rules with the first and second transformation requirements, it is possible that the cost of satisfying these requirements would be more than the cost of the third transformation requirement. In other words, satisfying the third transformation requirement could lead to a smaller data transformation than satisfying the first or second requirements. So for these rules the method should also do this comparison and select the transformation requirement with minimum cost. We consider all possible cases to achieve minimum data transformation. Finally, we have a database of α-discriminatory rules with their respective transformation requirements. An appropriate data transformation method (Phase 3) should be run to satisfy these requirements with minimum degree of information loss and maximum degree of discrimination removal.

4. DATA TRANSFORMATION METHOD

As mentioned above, 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 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 of Section V-B below. It is worth mentioning that decreasing the confidence of special rules (sensitive rules) by data transformation was previously used for knowledge hiding (Newman et al. 1998; Lewis, 1995; Thanh, 2011) in privacy-preserving data mining (PPDM). We assume that the class item C is a binary attribute. The details of our proposed data transformation method are summarized as follows:

1) For the α-discriminatory rules with the first transformation requirement (inequality $conf(X,Y \rightarrow CI) \leq conf(D,Y \rightarrow CI)/p$), the values of both sides of the inequality are independent, so the value of the left-hand side could be decreased without any impact on the value of the right-hand side. A possible solution for decreasing

$$conf(X,Y \rightarrow CI) = supp(X,Y,CI)/supp(X,Y) \qquad (1)$$

to any target value is to perturb the class item from CI to $\neg CI$ in the subset DBc of all records in the original dataset which completely support the rule $X,Y \rightarrow CI$ and have minimum impact on other rules to decrease the numerator of Expression (1) while keeping the denominator fixed. (Removing the records of the original dataset which completely support the rule $X,Y \rightarrow C$ would not help because it would decrease both the numerator and the denominator of Expression (1).)

2) For the α-discriminatory rules with the second transformation requirement (inequality $conf(X,Y \rightarrow D) \geq p$), the value of the right-hand side of the inequality is fixed so the value of the left-hand side could be increased independently. A possible solution for increasing

$$conf(X,Y \rightarrow D) = supp(X,Y,D)/supp(X,Y) \qquad (2)$$

above p is to perturb item D from $\neg D$ to D in the subset DYc of all records in the original dataset which completely support the rule $X,Y \rightarrow \neg D$ and have minimum impact on other rules to increase the numerator of Expression (2) while keeping the denominator fixed.

3) For the α-discriminatory rules with the third transformation requirement (inequality $conf(X,Y \rightarrow CI) < \alpha \cdot conf(Y \rightarrow CI)$), unlike the above cases, both inequality sides are dependent; hence, a transformation is required that decreases the lefthand side of the inequality without any impact on the right-hand side. A possible solution for decreasing

$$confXA,Y \rightarrow CI) = supp(X,Y,CI)/supp(X,Y) \qquad (3)$$

is to perturb item X from $\neg X$ to X in the subset DYc of all records of the original dataset which completely support the rule $\neg X,Y \rightarrow \neg CI$ and have minimum impact on other rules to increase the denominator of Expression (3) while keeping the numerator and $conf(Y \rightarrow CI)$ fixed. (Removing the records of the original dataset which completely support the rule $X,Y \rightarrow CI$ would not help because it would decrease both the numerator and the denominator of Expression (3) and also $conf(Y \rightarrow CI)$. Changing the class item CI would not help either because it would impact on $conf(Y \rightarrow CI)$. Records in DYc should be changed until the transformation requirement is met for each α-discriminatory rule. Among the records of DYc, one should change those with lowest impact on the other rules. Hence, for each record $dyc \in DYc$, the number of rules whose premise is supported by dbc is taken as the impact of*dyc*, that is $impact(dyc)$; the rationale is that changing dyc impacts on the confidence of those rules. Then the records dbc with minimum $impact(dyc)$ are selected for change, with the aim of scoring well in terms of the four utility measures proposed below. It means that transforming dyc with minimum $impact(dyc)$ 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.

5. UTILITY MEASURES

The proposed solution should be evaluated based on two aspects:

- The success of the proposed solution in removing all evidence of discrimination from the original dataset (degree of discrimination prevention).
- The impact of the proposed solution on data quality (degree of information loss).
- A discrimination prevention method should provide a good trade-off between both aspects above. The following measures are proposed for evaluating our 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 (DPP may not be 100% as a side-effect of the transformation process).
- *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 (side-effect of the transformation process).
- *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 (side-effect of the transformation process).

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 quality); ideally they should be 0%. MC and GC were previously proposed as information loss measures for knowledge hiding in PPDM (Luong, 2011).

6. DISCUSSION

Although there are some works about antidiscrimination in the literature, in this paper i introduced anti-discrimination for Recruiting Employees from Social Networks based on data mining. In this article problem statement (Saygin et al. 2001; Pedreschi et al. 2008; Verykios et al. 2004; Hajian et al. 2011), concentrated on discrimination discovery, by considering each rule individually for measuring discrimination without considering other rules or the relation between them. However in this work, we also take into account the PND rules and their relation with α-discriminatory rules in discrimination discovery. Then we propose a new preprocessing discrimination prevention method. In Section (Oliveira et al. 2006; Levine et al. 2008) also proposed a preprocessing discrimination prevention method. However, their works try to detect discrimination in the original data for only one discriminatory item based on a simple measure and then they transform data to remove discrimination. Their approach cannot guarantee that the transformed dataset is really discrimination-free, because it is known that discriminatory behaviors can often be hidden behind several items, and even behind combinations of them. Our discrimination prevention method takes into account several items and their combinations; moreover, we propose some measures to evaluate the transformed data in degree of discrimination and information loss.

7. CONCLUSIONS

I have examined how discrimination could impact on Recruiting Employees from Social Networks, especially IDS. IDS use computational intelligence technologies such as data mining. It is obvious that the training data of these systems could be discriminatory, which would cause them to make discriminatory decisions when predicting invasion. Our contribution concentrates on producing training data which are free or nearly free from discrimination while preserving their usefulness to detect real Discriminated Recruiting Employees from Social Networks. In order to control discrimination in a dataset, a first step consists in discovering whether there exists discrimination. If any discrimination is found, the dataset will be modified until discrimination is brought below a certain threshold or is entirely eliminated. In the future, we want to run our method on real datasets, improve our methods and also consider background knowledge (indirect discrimination).

SUMMARY OF RESEARCH

1. This article, composed within the limit of available resources, has provided useful information about discrimination discovery and prevention for recruiting employees from social networks and e-jobs.
2. It has availed scientists the opportunity to research more on the usefulness of Knowledge discovery about recruiting employees from online. e.g. social networks, online jobs websites.

FUTURE ISSUES

From the findings, recruiting peoples form internet is not feasible and also can find discrimination detection from peoples recruiting from on-campus (in-person).

DISCLOSURE STATEMENT

There is no financial support for this research work.

ACKNOWLEDGMENT

Much thanks to our guide for his constructive criticism, and assistance towards the successful completion of this research work.

REFERENCE

1. Hajian S, Domingo-Ferrer J, Mart´ınez-Balleste A. Rule protection for indirect discrimination prevention in data mining, in Proc.MDAI, LNCS 6820. Springer, 2011, 211–222
2. Levine R, Levkov A, Rubinstein Y. Racial discrimination and competition National Bureau of Economic Research, 2008
3. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 1995, 7, 57–78
4. Luong BT. Generalized discrimination discovery on semi-structured data supported by ontology. IMT Institute for Advanced Studies, Lucca, Italy, 2011, 14-18
5. Newman DJ, Hettich S, Blake CL, Merz CJ. UCI repository of Machine Learningdatabases, http://archive.ics.uci.edu/ml/datasets/Statlog+%28 German+Credit+Data%29, 1998.
6. Oliveira SRM, Zaiane OR. A unified framework for protecting sensitive association rules in business collaboration. *International Journal of Business Intelligence and Data Mining*, 2006, 24-25
7. Pedreschi D, Ruggieri S, Turini F. Discrimination-aware data mining. *Proc. of the 14th ACM International Conference on Knowledge Discovery and Data Mining (KDD 2008)*, ACM, 2012, 560- 568
8. Pedreschi D, Ruggieri S, Turini F. Discrimination-aware data mining. Proc. of the 14th ACM International Conference on Knowledge Discovery and Data Mining (KDD 2008), ACM, 2008, 560-568
9. Saygin Y, Verykios V, Clifton C. Using unknowns to prevent discovery of association rules. ACM SIGMOD Record, 2001, 45-54
10. Thanh BL. Generalized discrimination discovery on semistructured data, Ph.D. dissertation, IMT Institute for Advanced Studies, Lucca, Italy, 2011
11. Verykios V, Elmagarmid A, Bertino E, Saygin Y, Dasseni E. Association rule hiding. *IEEE Trans. on Knowledge and Data Engineering*, 16(4), 434-447, 2004