The Comparative Study of Data Mining Applications

Shirshendu Maitra, Pankaj Mudholkar

Assistant Professor

Department of Computer Science

Thakur Institute of Management Studies,

Career Development and Research (TIMSCDR)

Abstract—in this paper we have centered an assortment of strategies, methodologies and diverse regions of the exploration which are useful and stamped as the imperative field of information mining Technologies. As we know that numerous MNC's what's more, vast associations are worked in better places of the distinctive nations. Every spot of operation may create vast volumes of information. Corporate leaders require access from every single such source and take key choices .The information stockroom is utilized as a part of the huge business esteem by enhancing the adequacy of administrative choice making. In a questionable and profoundly aggressive business environment, the estimation of key data frameworks, for example, these are effortlessly perceived however in today's business surroundings, effectiveness or velocity is not by any means the only key for aggressiveness. This sort of gigantic measure of information's are accessible as tera to peta bytes which has definitely changed in the zones of science and building. To dissect, oversee and settle on a choice of such kind of colossal measure of information we require systems called the information mining which will changing in numerous fields. This paper bestows more number of uses of the information mining furthermore o centers extent of the information mining which will accommodating in the further research.

Keywords— *Information mining undertaking, Data mining life cycle , Visualization of the information mining model , Data mining Methods, Information mining applications*

I. INTRODUCTION

In the 21st century the people are utilized as a part of the diverse innovations to sufficient in the general public . Every last day the individuals are utilizing the unlimited information and these information are in the distinctive fields .It might be as archives, may be graphical configurations ,may be the video ,may be records (shifting exhibit) .As the information are accessible in the diverse organizations so that the best possible move to be made. To dissect these information as well as take a decent choice and keep up the information .As and when the client will required the information ought to be recovered from the database

and settle on the better choice .This method is really we called as an information mining or Knowledge Hub or just KDD(Knowledge Discovery Process).The imperative reason that pulled in a lot of consideration in data innovation the revelation of helpful data from extensive accumulations of information industry towards field of "Information mining" is expected to the impression of "we are information rich yet data poor". There is immense volume of information however we barely ready to turn them into valuable data and learning for administrative choice making in business. To create data it requires enormous accumulation of information. It might be distinctive arrangements like sound/video, numbers, content, figures, Hypertext group . To exploit information; the information recovery is just insufficient, it requires an instrument for programmed synopsis of information, extraction of the pith of data put away, and the revelation of examples in crude information. With the gigantic measure of information put away in records, databases, and different archives, it is progressively vital, to grow intense apparatus for examination and translation of such information and for the extraction of intriguing learning that could help in choice making. The main response to all above is 'Information Mining'. Information mining is the extraction of concealed prescient data from extensive databases; it is a capable innovation with extraordinary potential to offer associations some assistance with focusing on the most imperative data in their information stockrooms.

Information mining, prevalently known as Knowledge Discovery in Databases (KDD), it is the Nontrivial extraction of verifiable, already obscure and possibly valuable data from information in databases [3, 5]. It is really the procedure of discovering the concealed data/example of the archives.

THE DATAMINING TASK

The information mining errands are of diverse sorts relying upon the utilization   of information mining result the information mining errands are grouped as[1,2]

Exploratory Data Analysis

In the archives endless measure of data's are accessible .This information mining errand will serve the two purposes

1) Without the learning for what the client is looking, then
2) It investigate the information

These systems are intuitive and visual to the client.

Descriptive Modeling:

It depict all the information, it incorporates models for general likelihood conveyance of the information, apportioning of the p-dimensional space into gatherings and models depicting the connections between the variables.

Predictive Modeling:

This model permits the value of one variable to be predicted from the known values of other variables.

Discoving Patterns and Rules

This model allows the estimation of one variable to be anticipated from the known estimations of other variables. This undertaking is principally used to locate the concealed example and to find the example in the bunch. In a bunch various examples of distinctive size and groups are accessible.

The point of this undertaking is "the way best we will distinguish the examples" .This can be refined by utilizing standard incitement and numerous more strategies in the information mining calculation like(K-Means /K-Medoids) .These are known as the grouping calculation.

Retrieval by Content:

The essential goal of this errand is to discover the information sets of as often as possible utilized as a part of the for sound/video and in addition pictures It is discovering example like the example of enthusiasm for the information set.

TYPES OF DATA MINING SYSTEM

Information mining frameworks can be ordered by criteria the grouping is as follows [3]:

Classification of data mining systems according to the type of data source mined:

In an association a colossal measure of information's are accessible where we have to characterize these information however these are accessible a large portion of times in a comparative manner. we have to group these information as per its type(maybe sound/video ,content organization and so forth)

Classification of data mining systems according to the data model:

There are such a large number of number of information mining models(Relational information model, Object Model, Object Arranged information Model, Hierarchical information Model/W information model)are accessible and every last model we are utilizing the distinctive information .According to these information show the information mining framework group the information in the model.

Classification of data mining systems according to the kind of knowledge discovered:

This order in view of the sort of learning found or information mining functionalities, for example, portrayal, separation, affiliation, grouping, grouping, and so on. A few frameworks have a tendency to be extensive frameworks offering a few information mining functionalities together.

Classification of data mining systems according to mining techniques used:

This characterization is as indicated by the information examination methodology, utilized; for example, machine learning, neural systems, hereditary calculations, insights, perception, database arranged or information distribution center situated, and so on. The characterization can likewise consider the level of client cooperation included in the information mining process, for example, question driven frameworks, intelligent exploratory frameworks, or autonomous frameworks. A thorough framework would give a wide assortment of information mining systems to fit diverse circumstances and alternatives, and offer distinctive degrees of client communication.

VISUALIZING DATA MINING MODEL

The principle target of information representation is the general thought regarding the Information mining model .In information mining a large portion of the times we are recovering the information from the vaults which are in the covered up structure. This is the troublesome errand for a client. So this perception of the information mining model helps us to give most extreme levels of comprehension and trust.

The prescient model makes expectation about obscure information values by utilizing the known qualities. Ex. Grouping, Regression, Time arrangement examination, Prediction and so forth. The unmistakable model recognizes the examples or connections in information and investigates the properties of the information inspected. Ex. Bunching, Summarization, Association principle, Sequence disclosure and so forth.

A large portion of the information mining applications are planned to foresee the future condition of the information. Forecast is the procedure of breaking down the present and past conditions of the property and expectation of its future state. Order is a method of mapping the objective information to the predefined bunches or classes, this is a regulate learning in light of the fact that the classes are predefined before the examination of the objective information. The relapse includes the learning of capacity that guide information thing to genuine esteemed expectation variable. In the time arrangement examination the estimation of a trait is inspected as it differs after some time. In time arrangement examination is utilized for some measurable strategies which will break down the time-arrangement information, for example, auto relapse strategies etc. It is a few times utilized as a part of the two kind of displaying (1) ARIMA (II) Long-memory time-arrangement demonstrating

The term bunching means examines the diverse information objects without counseling a known class levels. It is likewise alluded to as unsupervised learning or division. It is the apportioning or division of the information into gatherings or bunches. The groups are characterized by concentrating on the conduct of the information by the area specialists. The term division is utilized as a part of particular connection; it is a procedure of apportioning of database into disjoint gathering of comparative tuples. Outline is the system of showing the outline data from the information. The affiliation principle finds the relationship between the diverse qualities. Affiliation standard mining is a two-stage process:

Discovering all successive thing sets, Generating solid affiliation rules from the continuous thing sets.

Succession revelation is a procedure of discovering the arrangement designs in information. This arrangement can be utilized to comprehend the pattern.

DATA MINING METHODS:

A portion of the prominent information mining routines are as per the following:

- Decision Trees and Rules
- Nonlinear Regression and Classification Methods
- Example-based Methods
- Probabilistic Graphical Dependency Models
- Relational Learning Models

We discovered these are some celebrated information mining systems are comprehensively named: On-Line Systematic Processing (OLAP), Classification, Clustering, Association Rule Mining, Temporal Information Mining, Time Series Analysis, Spatial Mining, Web Mining and so on. These strategies use distinctive sorts of calculations and information. The information source can be information stockroom, database level record or content document.

The calculations may be Statistical Algorithms, Decision Tree based, Nearest Neighbor, Neural Network based, Genetic Algorithms based, Ruled based, Support VectorMachine and so on. For the most part the information mining calculations are completely indigent of the two elements these are:

1) Which sort of information sets are utilizing?
2) What sort of necessities of the client?

Basing upon the above two variables the information mining calculations are used. A learning revelation (KD) procedure includes preprocessing information, picking an information mining calculation, and post handling the mining results. The Intelligent Discovery Assistants [7] (IDA), helps clients in applying substantial information disclosure forms. The IDA can furnish clients with three advantages:

1) A methodical list of substantial learning disclosure forms;

2) Effective rankings of legitimate procedures by diverse criteria, which pick between the alternatives;

3) A foundation for sharing learning, which prompts system externalities.

A few different endeavors have been made to computerize this procedure and outline of a summed up information mining device that gang's insight to choose the information and information mining calculations and up to some degree the learning revelation.

DATA MINING APPLICATIONS

Data Mining Applications in Healthcare

In any case, the achievement of human services information mining depends on the accessibility of clean social insurance information. In this admiration, it is important that the human services industry investigate how information can be better caught, put away, arranged and mined. Conceivable headings incorporate the institutionalization of clinical vocabulary and the sharing of information crosswise over associations to upgrade the advantages of human services information mining applications

Data mining is used for market basket analysis

Information mining procedure is utilized as a part of MBA (Market Basket Analysis).When the client need to purchasing a few items then this procedure offers us some assistance with finding the relationship between diverse things that the client put in their shopping containers. Here the disclosure of such affiliations that advances the business system .along these lines the retailers utilizes the information mining procedure so that they can recognize what clients intension (purchasing the distinctive pattern).In along these lines this method is utilized for benefits of the business furthermore buys the related things.

The data mining is used an emerging trends in the education system in the whole world

In Indian culture a large portion of the folks are uneducated .The fundamental point of in Indian government is the quality training not for amount. Be that as it may, the step by step the training frameworks are changed and in the 21st century an immense number of universalities are built up by the request of UGC. As the quantities of colleges are set up one next to the other, every last day a thousand years of understudies are selects the nation over. With immense number of advanced education hopefuls, we trust that information mining innovation can offer connecting learning some assistance with gapping in higher instructive frameworks. The shrouded examples, affiliations, and irregularities that are found by information mining methods from instructive information can enhance choice making procedures in higher instructive frameworks. This change can bring favorable circumstances, for example, boosting instructive framework proficiency, diminishing understudy's drop-out rate, and expanding understudy's advancement rate, expanding understudy's consistency standard in, expanding understudy's move rate, expanding instructive change proportion, expanding understudy's prosperity, expanding understudy's learning result, and decreasing the expense of framework forms. In this present time we are utilizing the KDD and the information digging instruments for removing the information this learning can be utilized for enhancing the nature of training .The choices tree order is utilized as a part of this kind of utilizations.

Data mining is now used in many different areas in manufacturing engineering

When we recover the information from assembling framework then the client is to utilize this information for diverse purposes like to discover the blunders in the information, to upgrade the configuration philosophy, to make the great nature of the information, how best the information can be bolstered for settling on the choice. Be that as it may the vast majority of time the information can be initially broke

down then after locate the concealed examples which will be control the assembling procedure which will further upgrade the nature of the items .Since the significance of information mining in assembling has unmistakably expanded in the course of the most recent 20 years, it is presently proper to basically survey its history and Application

Data Mining Applications can be generic or domain specific.

Information mining framework can be connected for nonexclusive or area particular. Some non specific information mining applications can't take its own these choices however manage clients for choice of information, determination of information digging technique and for the translation of the outcomes. The multi operators based information mining application [8, 10] has capacity of programmed choice of information mining procedure to be connected. The Multi Agent System utilized at diverse levels [8]: First, at the level of idea pecking order definition then at the outcome level to exhibit the best adjusted choice to the client. This choice is put away in learning Base to use in a later choice making. Multi Agent System Tool utilized for non specific information mining framework improvement [10] utilizes diverse specialists to perform distinctive errands.

A multi-tier data mining system is proposed to enhance the performance of the data mining process[9].

It has fundamental segments like client interface, information mining administrations, information access administrations and the information. There are three unique architectures displayed for the information mining framework to be specific one-tire, Two-tire and Three-tire building design. Non specific framework required to coordinate the same number of learning calculations as would be prudent and chooses the most suitable calculation to utilize. CORBA (Common Item Request Broker Architecture) has elements like: Integration of diverse applications coded in any programming dialect impressively simple. It permits reusability feasibly lastly it makes conceivable to manufacture extensive and versatile framework. The information mining framework construction modeling based on CORBA is given by Object Management Group [10] has all attributes to fulfill a conveyed and protest situated calculation. Information driven center and computerized techniques makes information mining open to no specialists [11]. The utilization of abnormal state interfaces can actualize the computerized techniques that shroud the information mining ideas far from the clients. Information driven configuration covers up away every one of the points of interest of mining philosophy and uncovered them through abnormal state errands that are objective arranged. These objective situated errands are executed utilizing information driven APIs. This outline makes information mining errand like different sorts of inquiries that clients perform on the information. In information mining better results could be gotten if huge information is accessible. It prompts the blending and connecting of neighborhood databases. Another information mining building design in light of Internet innovation tended to this issue. [12] The setting element plays basic part in the achievement of information mining. The significance and importance of same information in the diverse connection is distinctive. An information in one connection is imperative may not be much vital in other connection. Setting mindful information mining system channels helpful and intriguing connection variables, also, can deliver exact and exact expectation utilizing those factors [24].

Application of Data Mining techniques in CRM

Information mining system is utilized as a part of CRM .Now a days it is one of the hotly debated issues to investigate in the industry in light of the fact that CRM have pulled in both the experts and scholastics. It means to give research rundown on the utilization of information mining in the CRM area and methods which are regularly utilized. In spite of the fact that this audit can't claim to be comprehensive, it provides

sensible bits of knowledge and demonstrates the rate of examination on this subject. The outcomes exhibited in this paper have a few critical ramifications: Research on the utilization of information mining in CRM will increment essentially later on in light of past production rates and the expanding enthusiasm for the zone. Most of the checked on articles identify with client maintenance [19]

The Domain Specific Applications

The area particular applications are engaged to utilize the space particular information and information digging calculation that focused for particular target. The applications concentrated on in this setting are planned to create the particular information. In the diverse areas the information creating sources produce diverse sort of information. Information can be from a basic content, numbers to more intricate sound video information. To mine the examples and in this way learning from this information, distinctive sorts of information mining calculations are utilized. The gathering and choice of connection particular information and applying the information mining calculation to produce the connection particular information is in this way an apt occupation. In numerous areas particular information mining applications the space specialists assumes fundamental part to mine valuable information.

In the distinguishing proof of outside highlighted French the sound documents were utilized and the best 20 information mining calculations were applied[13] the Logistic Regression model discovered the most hearty calculation than other calculation.

In language research and language

Building much time additional etymological data is required around content. A phonetic profile that contains substantial number of semantic components can be produced from content record consequently utilizing information mining [14]. This procedure discovered entirely compelling for initiation check and acknowledgment. A profiling framework utilizing blend of lexical and syntactic components demonstrates 97% exactness in selecting right creator for the content. The phonetic profiling of content successfully used to control the nature of dialect and for the programmed dialect verification.[15] This technique confirms consequently the content is of local quality. The results demonstrate that dialect check is in fact conceivable.

In Medical Science

In restorative science there is substantial degree for use of information mining. Finding of diesis, wellbeing care, patient profiling and history era and so on are the couple of samples. Mammography is the system utilized as a part of bosom tumor identification. Radiologists face parcel of challenges in recognition of tumors that is the reason CAM (Computer Aided Methods) could serves to the therapeutic staff .With the goal that they can produce the great nature of the outcome discovery [16]. The neural systems with back-engendering what's more, affiliation guideline digging utilized for tumor arrangement as a part of mammograms. The information mining successfully utilized as a part of the analysis of lung variation from the norm that may be malignant or kindhearted [17]. The information mining calculations fundamentally diminish quiet's dangers and conclusion costs. Utilizing the forecast calculations the watched expectation precision was 100% for 91.3% cases. The utilization of information mining in medicinal services is the broadly utilized utilization of information mining. The therapeutic information is perplexing and hard to break down. A REMIND (Reliable Extraction and Meaningful Inference from Non-organized Data) framework [21] coordinates the organized and unstructured clinical information in patient records to consequently make excellent organized clinical information. To receive the high quality

system, we can mine the current patient records to bolster rules and give consistence to enhance patient consideration. [21].

Data Mining methods are used in the Web Education

Information mining systems are utilized as a part of the web Education which is utilized to enhance courseware. The connections are found among the utilization information got amid understudies' sessions. This learning is extremely valuable for the instructor or the creator of the course, who could choose what adjustments, will be the most fitting to enhance the viability of the course. [22].In the 21st century the novices are utilizing the information mining systems which are one of the best learning techniques in this era [21]. This makes it conceivable to build the familiarity with learners. Web Education which will quickly development in the utilization of information mining techniques to instructive visits, which is both achievable and can be change in learning situations in the 21st century.

CONCLUSION

In this paper we quickly checked on the different information mining applications. This audit would be supportive to specialists to concentrate on the different issues of information mining. In future course, we will audit the different characterization calculations and noteworthiness of transformative figuring (hereditary programming) methodology in outlining of proficient order calculations for information mining. The vast majority of the past studies on information mining applications in different fields utilize the assortment of information sorts range from content to pictures and stores in assortment of databases and information structures. The diverse routines for information mining are utilized to remove the examples and accordingly the learning from this assortment databases. Determination of information and techniques for information mining is a vital errand in this process and needs the learning of the area. A few endeavors have been made to outline and build up the non specific information mining framework however no framework discovered totally non specific. Therefore, for each space the area master's right hand is obligatory. The area specialists should be guided by the framework to successfully apply their insight for the utilization of information mining frameworks to produce required information. The area specialists are required to decide the assortment of information that ought to be gathered in the particular issue space, determination of particular information for information mining, cleaning what's more, change of information, removing examples for learning era lastly understanding of the examples and learning era. The majority of the area particular information mining applications show exactness above 90%. The non specific information mining applications are having the confinements. From the investigation of different information mining applications it is watched that, no application called nonspecific application is 100 % non specific. The astute interfaces and insightful operators up to some degree make the application non specific yet have confinements. The area specialists assume critical part in the diverse phases of information mining. The choices at distinctive stages are affected by the elements like space and information points of interest, point of the information mining, and the connection parameters. The space particular applications are expected to concentrate particular information. The area specialists by considering the client's prerequisites and other connection parameters control the framework. The outcomes yield from the area particular applications is more precise and valuable. Accordingly it is reason that the area particular applications are more particular for information mining. From above study it appears to be exceptionally hard to outline and build up an information mining framework, which can work progressively for any space.

REFERENCES

[1] Introduction to Data Mining and Knowledge Discovery , Third Edition ISBN: 1-892095-02-5, Two Crows Corporation, 10500 Falls Road, Potomac, MD 20854 (U.S.A.), 1999.

[2] Larose, D. T., "Discovering Knowledge in Data: An Introduction to Data Mining", ISBN 0-471-66657-2, john Wiley & Sons, Inc, 2005.

[3] Dunham, M. H., Sridhar S., "Data Mining: Introductory and Advanced Topics", Pearson Education, New Delhi, ISBN: 81-7758-785-4, 1stEdition, 2006

[4] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Rei nartz, T., Shearer, C. and Wirth, R... "CRISP-DM 1.0: Step-by-step data mining guide, NCR Systems Engin eering Copenhagen (USA and Denmark),DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringenen Bank Group B.V (The Netherlands), 2000".

[5] Fayyad, U., Piatetsky-Shapiro, G., and Smyth P., "F rom Data Mining to Knowledge Discovery in Databases," AI Magazine, American Association for Artificial Intelligence, 1996.

[6]Tan Pang-Ning, Steinbach, M., Vipin Kumar. "Introduction to Data Mining", Pearson Education, New International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.3, June 2012 56 Delhi, ISBN: 978-81-317-1472-0, 3rd Edition, 2009. Bernstein, A. and Provost, F., "An Intelligent Assistant for the Knowledge Discovery Process", Working Paper of the Center for Digital Economy Research, New York University and also presented at the IJCAI 2001 Workshop on Wrappers for Performance Enhancement in Knowledge Discovery in Databases.

[7] Baazaoui, Z., H., Faiz, S., and Ben Ghezala, H., "A F ramework for Data Mining Based Multi-Agent: An Application to Spatial Data, volume 5, ISSN 1307-6884," Proceedings of World Academy of Science, Engineering and Technology, April 2005.

[8] Rantzau, R. and Schwarz, H., "A Multi-Tier Architec ture for High-Performance Data Mining,A Technical Project Report of ESPRIT project, The consortium of CRITIKAL project, Attar Software Ltd. (UK), Gehe AG (Denmark); Lloyds TSB Group (UK), Parallel Applications Centre, University of Southampton (UK), BWI, University of Stuttgart (Denmark), IPVR, University of Stuttgart Denmark)".

[9] Botia, J. A., Garijo, M. y Velasco, J. R., Skarmeta, A. F., "A Generic Data mining System basic design and implementation guidelines", A Technical Project Report of CYCYT project of Spanish Government.1998.WebSite: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.53.1935

[10] Campos, M. M., Stengard, P. J., Boriana, L. M., "Da ta-Centric Automated Data Mining", , Web Site.: www.oracle.com/technology/products/bi/odm/pdf/automat ed_data_mining_paper_1205.pdf

[11] Sirgo, J., Lopez, A., Janez, R., Blanco, R., Abajo, N., T arrio, M., Perez, R., "A Data Mining Engine based on Internet, Emerging Technologies and Factory Automation," Proceedings ETFA '03, IEEEz Conference,16-19Sept.2003.WebSite:www.citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.8955

[12] Bianca V. D.,Philippe Boula de Mareüil and Martine A dda-Decker, "Identification of foreign-accented French using data mining techniques, Computer Sciences Laboratory for Mechanics and Engineering Sciences (LIMSI)".Website www.limsi.fr/Individu/bianca/article/Vieru&Boula&Madda_ParaLing07.pdf

[13] Bianca V. D.,Philippe Boula de Mareüil and Martine A dda-Decker, "Identification of foreign-accented French using data mining techniques, Computer Sciences Laboratory for Mechanics and Engineering Sciences (LIMSI)".Website www.limsi.fr/Individu/bianca/article/Vieru&Boula&Madda_ParaLing07.pdf

[14] Halteren, H. V., Oostdijk N., "Linguistic profiling of texts for the purpose of language verification, The ILK research group, Tilburg centre for Creative Computing and the Department of Communication and Information Sciences of the Faculty of Humanities, Tilburg University, The Netherlands."Website:www.ilk.uvt.nl/~antalb/textmining/LingProfColingDef.pdf

[15]Antonie, M. L., Zaiane, O. R.,Coman, A., "Application of Data Mining Techniques for Medical Image Classification", Proceedings of the Second International Workshop on Multimedia Data Mining MDM/KDD 2001) in conjunction with ACM SIGKDD conference, San Francisco, August 26, 2001.

[17] Kusiak, A., Kernstine, K.H., Kern, J.A., McLaughlin, K.A., and Tseng, T.L., "Data Mining: Medical And Engineering Case Studies". Proceedings of the Industrial Engineering Research 2000 Conference, Cleveland, Ohio, pp. 1-7,May 21-23, 2000.

[18] Luis, R., Redol, J., Simoes, D., Horta, N., "Data Warehousing and Data Mining System Applied to E-Learning, Proceedings of the II International Conference on Multimedia and Information & Communication Technologies in Education, Badajoz,Spain, December 3-6th 2003.

[19] Chen, H., Chung, W., Qin, Y., Chau, M., Xu, J. J., Wang,G., Zheng, R., Atabakhsh, H., Crime Data Mining: An Overview and Case Studies", A project under NSF Digital Government Programme, USA, "COPLINK Center: Information and Knowledge Management for Law Enforcement,", July 2000 -June 2003

[20] Kay, J., Maisonneuve, N., Yacef, K., Zaiane O., "Mining patterns of events in students' teamwork data", Proceedings of the ITS (Intelligent Tutoring Systems) 2006 Workshop on Educational Data Mining, pages 45-52, Jhongli, Taiwan, 2006.

[21]Rao, R. B., Krishnan, S. and Niculescu, R. S., "Data Mining for Improved Cardiac Care" , SIGKDD Explorations Volume 8, Issue 1.

[22]Ghani, R., Probst, K., Liu, Y., Krema, M., Fano, A., "Text ining for Product Attribute Extraction", SIGKDD Explorations Volume 8, Issue 1.

[23] DeBarr, D., Eyler-Walker, Z., "Closing the Gap: Automated Screening of Tax Returns to Identify Egregious Tax Shelters". SIGKDD Explorations Volume 8, Issue 1.

[24]Kanellopoulos, Y., Dimopulos, T., Tjortjis, C., Makris, C. "Mining Source Code Elements for Comprehending Object-Oriented Systems and Evaluating Their Maintainability", SIGKDD Explorations Volume 8, Issue 1.