A Survey on Orphanages Circumstances Using Data Mining

[1]Surinder Kaur, [2]Kanwalpreet Singh Attwal

[1]Student, Department of Computer Engineering, Punjabi University, India

[2]Department of Computer Engineering, Punjabi University, India

Abstract: **Data Mining is a technique used in various domains to give meaning to the available data**. **Data mining is the most important part of computers because everything is depended on the data today. So to maintain this huge data it is stored in databases & it is very difficult to search out particular information from huge database. Data mining helps here to mine the particular information from any of the databases. In data mining terms, a record in a child table without an associated record in parent table is known as an orphan. Data mining helps here to analysis the records of orphans, their needs &rights. This paper presents the orphanages lifestyle closely and discusses the strartegy to be followed to bring some changes in their lifestyle.**

Keywords: **Data mining, Data Warehouse, KDD, orphans, orphanages.**

I.   INTRODUCTION

Data mining refers to extracting or "mining" knowledge from large amounts of data. It is information or patterns from large information repositories such as relational databases, data warehouses (A data warehouse is a repository of information collected from multiple sources, stored under a unified schema, and usually residing at a single site) etc. The goal of the data mining process is to extract information from a data set and transform it into an understandable structure for future use. Data Mining, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data Mining is about solving problems by analyzing data already present in databases [1]. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. Data mining tools can answer business questions that traditionally were too time consuming to resolve. It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data pattern analysis [2]. The unifying goal of the KDD process is to extract knowledge from data in the context of large databases. Knowledge discovery as a process consists of an iterative sequence of the following steps[3]:

1. Data cleaning (to remove noise and inconsistent data)

2. Data integration (where multiple data sources may be combined)

3. Data selection (where data relevant to the analysis task are retrieved from the database)

4. Data transformation (where data are transformed into forms appropriate for mining)

5. Data mining (an essential process where intelligent methods are applied in order to extract data patterns)

6. Pattern evaluation (to identify the truly interesting patterns representing knowledge based on some interestingness measures)

7. Knowledge presentation (where visualization and knowledge representation techniques are used to present the mined knowledge to the user),as shown in fig. 1.
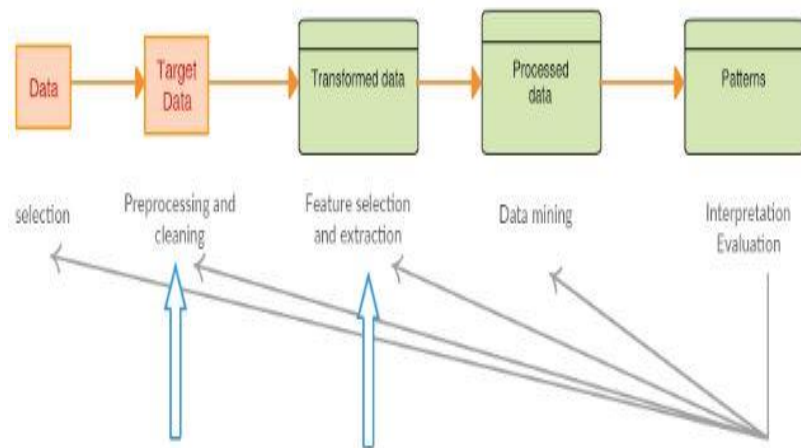
Fig. 1 A KDD Process

Data mining techniques:-

There are several major *data mining techniques that* have been developed and used in data mining projects

A. Association:

Association is one of the best known data mining technique. In association, a pattern is discovered based on a relationship of a particular item on other items in the same transaction. For example, the association technique is used in *market basket analysis* to identify what products that customers frequently purchase together. Based on this data businesses can have corresponding marketing campaign to sell more products to make more profit.

B. Classification:

Classification is a classic data mining technique based on machine learning. Basically classification is used to classify each item in a set of data into one of predefined set of classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics. Classification is a form of data analysis that extracts models describing important data classes[3].

Classification consists of predicting a certain outcome based on a given input. In order to predict the outcome, the algorithm processes a training set containing a set of attributes and the respective outcome, usually called goal or prediction attribute. The algorithm tries to discover relationships between the attributes that would make it possible to predict the outcome. Next the algorithm is given a data set not seen before, called prediction set, which contains the same set of attributes, except for the prediction attribute – not yet known.

C. Clustering:

Clustering is a data mining technique that makes meaningful or useful cluster of objects that have similar characteristic using automatic technique. Different from classification, clustering technique also defines the classes and put objects in them, while in classification objects are assigned into predefined classes. To make the concept clearer, we can take library as an example. In a library, books have a wide range of topics available. The challenge is how to keep those books in a way that readers can take several books in a specific topic without hassle. By using clustering technique, we can keep books that have some kind of similarities in one cluster or one shelf and label it with a meaningful name. If readers want to grab books in a topic, he or she would only go to that shelf instead of looking the whole in the whole library.

D. Prediction:

The prediction as it name implied is one of a data mining techniques that discovers relationship between independent variables and relationship between dependent and independent variables. For instance, prediction analysis technique can be used in sale to predict profit for the future if we consider sale is an independent variable, profit could be a dependent variable. Then based on the historical sale and profit data, we can draw a fitted regression curve that is used for profit prediction.

II.   INTRODUCTION TO ORPHANS & ORPHANAGES

An orphan can be defined as a child who lost one parent or both parents or whose parents are dead or have abandoned them permanently. With the help of data storage policy it is easy to investigate the orphan records to find out where they belonged. According to government policies everybody on the earth have their own unique id, in the same way orphans have their own unique records in terms of government policy and all these birth date, name etc are recorded in their orphanages in their records. Earlier these records were manually created and lots of files have been saved in the records and it becomes very time consuming to find out the particular orphan records .To make this task easy database helps here to maintain their records via rows and columns. Data mining is very powerful concept of extracting any kind of knowledge in few seconds. Databases help here to updating, deleting, inserting any kind of data in records with using just id of the orphan without disturbing the other records of orphans.

In our country there are lots of orphans. Orphans are worried to stay alive and as an outcome sometimes they live very dangerous routine as street Orphans, beggars, criminal, or child prostitutes. They don't know about themselves e.g.name, parents name, home address, school, birthday also because in their life they rarely live with their family so how they know about anything .For these kind of children government mentioned some policy e.g. in every orphanage this is necessary to record every orphan's in & out dates . It is recorded as an orphan take admission in orphanage on that particular date & time. This is recorded in the orphanages records e.g. databases. This is important for population counting & also for allowing of budget to every orphanage. Databases helps here to maintain records of every orphans who lives in the orphanage from when to when & it's registering from first day to adoption dates are recorded. With the help of data mining it is quite easy to find out any of the orphan from any of the orphanage. Governments usually have policies that necessitate organizations to search for approval to set up housing care for Orphans and to record such facilities. It is quite important about the staffing structure in the orphanages because everything in the orphanages is directly reflected upon the life of the orphans. If the staff is caring and they have ability to take responsibility of the orphans then they can handle them otherwise orphans can never be survived with this alone life without care and love of their caretakers. This is necessary to record all the data of the staff in government records because a number of orphans are to live in the hands of this staff.

Data mining helps to quick retrieval of information about the orphanages. It is time saving because by this orphanages can easily search orphans records through internet & it also provides the awareness among people with these websites.

III.   LITERATURE SURVEY

T.V V Satya Surya Chandra Divya1, V.Sireesha 2, S.Krishna Chaitanya 3, N. Baswanth 4 (2015) This paper is based on the real time project work that is helpful in knowing the details about the orphanages around the city. Still people have very lack of awareness about places where the orphanages are & orphans live. As same the donators won't find out a proper way to donate. In these cases the records are build. The members of orphanage call them time to time for asking the donation. Data mining helps here to maintain & extracting the records. In the datasets, the Administration is the only person who canenter the details of the Orphan with a specific authentication & the user only view the information about Orphans & donations.

*Nasir Mohammad Bhat (2014)[6]*A Study of Emotional Stability and Depression in Orphan Secondary School Students: The death of a parent is a risk factor for the development of psycholosocial issues in children. The paper carries out the study of mental health status among orphan secondary school students in comparison with non orphan secondary school students of Kasmir Valley. Purposive sampling technique was used for research purpose. For the assessment of Emotional stability among orphan and non orphan adolescents, mental health scale developed by Alpana Sen Gupta and Arun Kumar singh (1983) was used. For depression Beck's Depression Inventory was used. To study the difference between the groups t-test was used. Result findings suggest significant differences in the Emotional stability and Depressive levels between the two groups. Orphans were found at the lower side of Emotional stability and higher levels of depression as compared to non orphans secondary school students.

Sudha Katyal (2015)[9] The objective of this study was to investigate resilience among orphan and non-orphan children. The study was carried out in two orphanages and two schools run by Government in Tricity- Chandigarh, Panchkula and Ajitgarh (Mohali) in India. The sample consisted of 50 orphan children (25 from each orphanage) and 50 non-orphan children (25 from each school) aged 12-18 years. Standardized tool was used to assess resilience of the subjects. The

findings revealed that major reason for living in the orphanage was death of parents, closely followed by financial problems of single parenthood. Majority of children were left in these institutes by their mothers and relatives. Most of orphan children had relatives too, whom they visited 1-2 times in a year. There was significant difference in resilience of orphan and non-orphan children, with orphan children having higher resilience than that of non-orphan children. For analyzing the data, percentage, mean, standard deviation, standard error of mean and 't' test were used.

Ombuya, Benjamin Dickens, Yambo, John M. Onyango, Omolo Tom Mboya (2012)[13] Effects of Orphan hood on Girl-child's Access and Retention in Secondary School Education: A Case of Rongo District, Kenya depict this study focused on orphan hood which is a critical issue in almost every household not only in Kenya but also the world over. The objective of this article was; to identify the school related effects of orphan hood on the girl-child's access and retention in secondary school education in Rongo District, Kenya.

Simple random sampling was used to select 468 orphaned girls, and purposive random sampling was used to include the 36 head teachers and 36 teacher counselors totaling to 540 respondents. Data was collected using closed and open ended questionnaires. The questionnaires were pre-tested by being administered to 30 respondents from two schools which were not included in the study. The scores obtained from the first and second tests were 0.76 and 0.77 respectively. The descriptive statistics including percentages mean and frequency tables were used to analyze the data. Inferential statistics such as chi-square ($\chi2$) was also used to analyze quantitative data. All hypotheses testing procedures were based on $\alpha = 0.05$ level of significance. Data was analyzed using the Statistical Package for Social Sciences (SPSS) version 11.0 computer programme. The study found out that orphan hood significantly affects the girl-child's access and retention in secondary school education in Rongo District. The study recommends that guidance and counseling interventions be put in place to address school related, socio-cultural, socio-economic effects and needs of orphaned girl-child. The results of the study would help the government in formulation of policy and programme interventions aimed at helping the orphaned girl-child and to specifically address the needs of the families with such children by making the said policies and programmes as proactive as possible to the needs of the orphans.

Mira Saxe-Smith (2015)[19] The Lucrative Business of Children: A Study of Child Trafficking in Nepalese Orphanages, The purpose of this research paper is to shed light on and explore the detrimental effects orphanage volunteer tourism has on communities in Nepal, especially the orphans themselves and their families who become subject to different degrees of manipulation, deceit and abuse due to the industry. This study will also delve into the Nepal Civil War's effect on orphan trafficking in Nepal, the effects of institutionalization, media and campaigns launched against voluntourism, and other related topics. The paper particularly focuses on orphanage voluntourism in the tourist heavy and capital of Nepal, Kathmandu.

IV. PROBLEM FORMULATION

Data mining works in many applications, so I tried to analyze the work in social activity using data mining techniques and tools. At the adoption time person who wants to adopt a child, he/she has to give written application to the orphanage society .For this they have to give their copy of address proof, salary slips, passport, marriage, medical, police & notary certificate. These all are recorded in the databases of orphanage society. These records are well organized & if anybody needs to check any of the orphan knowledge from any place, data mining helps here to mine out the information according to the need.

In earlier days, people were not having much awareness about the orphanages. So this technology ideas, make people aware about this & data mining helps to search out the orphans via internet that will make easier for donor who wants to donate in orphanages. Databases maintain all the data about each & every orphan & orphanage .So it becomes very easy for people to search out the near orphanages to help them & people who looked for adoptions.

V. FUTURE WORK

This research relates to orphan needs, living requirements, survival, expectations and orphanage structures in terms of life ,living style, education, health, enjoyable terms, finances, and staffing using data mining which helps to analysis all the tasks of orphans .

Collecting the data via asking some questions to children who lives in home and who survive on roads about their living style and same questions would be asked from donor and members of the orphanages about the orphans needs and

changes they want in their life style. Data mining helps here to find out the results in terms which are easily understood. The results obtained would be analyzed by data mining tools and techniques.

It is a try to make future changes of orphans. This study would not only cares for the Orphan's basic needs but also helps to prepare them to be independent when the time comes. This is a try to make them notice in the world how they survive & what are their needs in life to live a normal life like other children who live with their parents.

REFERENCES

[1] Joseph, Zernik, "*Data Mining as a Civic Duty* – Online Public Prisoners Registration Systems", International Journal on Social Media: Monitoring, Measurement, Mining, vol. - 1, no.-1, pp. 84-96, September2010.

[2] Clustering And Classification: Data Mining Approaches by Ed Cole

[3] Data Mining Concepts And Techniques By Jiawei Han and Micheline Kamber

[4] Bathina, J. (2013). Educating to save the world. International Journal of Orphans's Spirituality, 18(2), 173–187.

[5] Wood, L., & Goba, L. (2011). *Care and support of orphaned and vulnerable Orphans at school*: Helping teachers to respond, 31, 275–290.

[6] Agarwal, K.N., Agarwal, D. N., Upadhyay, S. K., & Singh, S, "*Learning disability in rural primary schoolchildren*" Indian Journal of Medical Research, vol.94,pp .89-95,May 1991

[7] Ames, E. W," *The development of Romanian Orphanage Children Adopted to Canada*. Final Report to Human Resources Development, Canada and children's adjustment", American Psychologist, vol. 53, pp.167-184,Feb 1997

[8] Atwine , B ., Cantor-Grace, E., & Bajunirwe, F" *Psychological distress among autonomy in middle and late adolescence*: Connectedness, separation, attachment. Development and Psychopathology", vol. 3,pp.793-809.

[9] London: Routledge. 2 Cyrulnik, B. (2010).Early Attachment and the building of resilience: The theoretical basis of bereavement counselling. Retrieved on November 13, 2014 from http://www.sagepub.com/ Working-with-Bereaved Children-and-Young People.

[10] Nagar, D. (1992). Deprived Children. Jaipur, India: Printwell.

[11] Kiirya, S. (2005).Parental death and its effect on orphaned children's self-esteem and sociability at school. Ph.D Dissertation, Makerere University, Uganda.

[12] Thomas, D., Cas, G. A., Frankenberg, E., &Suriastini, W. (2011).The Impact of ParentaDeath on Child Wellbeing, 51(2), 437–457

[13] Ombuya, Benjamin Dickens, Yambo, John M. Onyango, Omolo Tom Mboya (2012) *Effects of Orphan hood on Girl-child's Access and Retention in Secondary School Education*

[14] Admassie, A. (2003) Child Labour and Schooling in the Context fa Subsistence rural economy: Can they be compatible? International Journal of Educational Development, 23(2):167-185.

[15] Ainsworth, M., Beegle, K, and Koda, G.(2005) *The Impact of Adult Mortality and Parental deaths on primary schooling in North-West Tanzania.* The Journal of Development Studies, 41(3):412-439.

[16] Ainsworth, M. and Filmer, D. (2006). Inequality in children's schooling: AIDS, Orphanhood, Poverty and Gender. World Development 34 (6): 1099 – 1128.

[17] Al Sammarai, S. and Peasgood, T. (1998)Educational Attainments and Household Characteristics in Tanzania. Economics of Review, 17(4): 395-417.

[18] Ames, P. (2004) Schooling for Girls in Rural Peru.Id21 (internet). Available from ext/insightseduzart7. html

[19] Danekhu, Ramesh. "*Stop the Orphanage Business*." Accessed May 6, 2015. "Don't Create More Orphans." Accessed April 20th, 2015.