International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) e-ISSN: 2395-5325
Volume3, Issue 2 (February '2017)

ANALYZING IMPACTS OF INDIANS EMPLOYING IN FOREIGN COUNTRIES USING DATA MINING APPROACHES

M. Priya,
PG Student,
Department of Computer Science,
Jayaraj Annapackiam College for Women (Autonomous),
Periyakulam, TamilNadu, India.

T. Kokila,
Assistant Professor,
Department of Computer Science,
Jayaraj Annapackiam College for Women (Autonomous),
Periyakulam, TamilNadu, India.

S. Banu,
PG Student,
Department of Computer Science,
Jayaraj Annapackiam College for Women (Autonomous),
Periyakulam, TamilNadu, India.

Abstract: The issue of foreign workers has received increase media and national attention. However, to date there has been limited research on the nature and consequences of employment of foreign workers. Introduction of significant changes in recruitment phenomenon has ended in painful and traumatic atmosphere which barely acceptable by local workforce in Foreign countries. This conceptual paper can be derived from the field of industrial relations which play a significant role in employment of foreign workers. This case study reports on the preliminary findings on employment status of local workforce and trade union rights affected by employment of foreign workers. In addition, the research makes a number of recommendations, including the need for further development on reducing the employment of foreign workers and more refined targeting of vulnerable foreign workers linked with labor legislations.

Keywords: Association Rule Mining Algorithm, Apriori Algorithm, Cluster Analysis, Data Mining.

I. INTRODUCTION

For over the past decade, it has over dependence on the foreign workers the number of foreign workers has increased. The economic growth of the country, in particular by alleviating labor shortages in selected sectors of the economy but they still bring a lot of disadvantages to the country in terms of employment opportunity for local workers and effectiveness roles of trade union movement.

According to the Economic Report 2010/2011 by the Finance Ministry, there were 1.8 million registered foreign workers in Malaysia, 38.2% were employed in the manufacturing sector, 16% in the construction and 14.2% in the plantation sectors. Indonesia accounted for the highest number of registered foreign workers in Malaysia at 50.9% followed by Bangladesh was second highest, accounting for 17% of the total foreign workers in Malaysia, Nepal at 9.7%, Myanmar, 7.8%, India, 6.3% and Vietnam, 4.2%.

Currently government has agreed to approve the recruitment of 45,000 foreign workers from India to meet the demand in 13 small-scale business sectors, which are facing manpower shortage in year 2011.

Significance of the study: Most of the people in America tend to work only 8 hours a day. Typically, they work from 8 am to 4.30 pm early. It is Common to take 30 min lunch break. Some eat at desk and some just grab a quick lunch. When it comes to importance of work, people consider work as just work not LIFE. It is just way the culture is, Work is just part of the life and not life. Unlike in India as I talk to my cousins, friends, juniors working in IT, they go to work at 9 AM or so and come home at 9 PM. Most of them stay at work for 12 hrs.

Objectives of the study: The objective of the study is to investigate the Impacts of Indians Employing in foreign countries. This is a case study of magnitude and listing the factors which influence the people to move towards foreign countries for their job.

- To analyze the most common reasons for people moving to foreign countries
- To mention the situation of their family

Null Hypothesis: The hypotheses that were framed, which are required to be tested by collecting adequate primary data in the future research, are as follows:

- One may either do not feel overseas.
- No friends, no family, no social life and all such things should be taken into notice before hand and should be prepared emotionally before leaving for abroad.

You never know what kinds of people you will come across.

Limitations: The study is based on impacts of Indians employing in foreign countries. Common factors are improving their status, addicted to foreign culture, good salary for secured future etc. Working abroad is working away from your homeland. Now a day it has become a

fashion to move abroad and work in order to earn money. People move from one country to another in search of job or to settle.

II. DATA ANALYSIS

2.1 Algorithm used

Data mining is the core process of knowledge discovery in database. It is the process of extraction of useful patterns from the large database. To analyze the large amount of collected information, the area of Knowledge Discovery in Database (KDD) provides techniques which extract interesting patterns in a reasonable amount of time .data mining is the application of efficient algorithms to detect the desired patterns contained within the given data. Data mining is the extraction of hidden descriptive or predictive information from large databases.

Association Rule Mining: Association rules mining are one of the major techniques of data mining. The purpose of association analysis is to figure out the hidden association and some useful rules of data base, and uses these rules to speculate and judge the unknown matter from the already known information. Association rule mining has many important applications in our life.

Association Rule : An association rule is one of the forms x=>y. and each rule has two basic needs: support and confidence. Things that occur often together can be associated to each other. These together occurring things form a frequent itemset. Conclusions based on the frequent itemsets make association rules.

2.2 Apriori Algorithm

Apriori algorithm is a fundamental algorithm mining association rule. It contains two processes:

- Detect all frequent itemsets by scanning db.
- Form strong association rules in the frequent itemsets.

Process one needs to scan DB several times, which consumes a lot of time and space. As a result, what needs to be improved is the mining competency of frequent group of things in DB. Apriori algorithm is a significant algorithm for mining frequent itemsets for Boolean association rules. Apriori algorithm is formed by Agrawal and Srikantin 1994. It is the most fundamental and important algorithm for mining frequent itemsets. Apriori is used to detect all frequent itemsets in a provided database db.Thekeynote of Apriori algorithm is to form multiple passes over the database. It employs an repetitive approach called as a breadth-first search (level-wise search).

2.2.1 Key Concepts

• **Frequent Itemsets:**The itemsets which has minimum help (denoted by li for i^{th}-itemsets), Apriori property: any subgroup of frequent things must be frequent.

• **Join Operation:** to detect lk, a group of candidate k-group of things is developed by adding lk-1with itself.

How Apriori Works?

- **Find All Frequent Itemsets.**
- **Get Frequent Things**: Things whose occurrence in database is more than or equal to the minimum help threshold.
- **Frequent Itemsets**: Develop candidates from frequent things. Prune the results to detect the frequent itemsets.Develop strong association rules from frequent itemsets.Rules which satisfy the minimum support and minimum confidence threshold**.**
- **Association Rule**:
  Association rule of data mining involves picking out the unknown inter-dependence of the data and finding out the rules between those items [3]. Agrawal introduced association rules for point of sale (POS) systems in supermarkets. A rule is defined as an implication of the form A=>B, where A∩ B≠Ø. The left-hand side of the rule is called as antecedent. The right-hand side of the rule is called as consequent.
- **Support:** I = { i_1,i_2,i_3, … , i_m} is a collection of items. T be a collection of transactions associated with the items. Every transaction has an identifier TID [6]. Association rule A=>B is such that A Є I, B Є I. A is called as Premise and B is called as Conclusion. The support ,S, is defined as the proportion of transactions in the data set which contains the itemset.
  Support(X=>Y) = Support (XUY) = P(XUY).
- **Confidence:** The confidence is defined as a conditional probability
  Confidence (X=>Y) = Support (XUY) / Support(X) = P(Y/X). Lift: is the ratio of the probability that L and R occur together to the multiple of the two individual probabilities for L and R, i.e. lift = Pr(L,R) / Pr(L).Pr(R).
- **Conviction:** is similar to lift, but it measures the effect of the right-hand-side not being true. It also inverts the ratio. So, a conviction is measured as:
  Conviction = Pr(L).Pr(not R) / Pr(L,R)

2.3 Sample Used

Association Rule mining algorithm in R

APRIORI is a level-wise breadth-first algorithm which counts transactions to find frequent itemsets and then derive association rules from them I apriori() in package arules.

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 50

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[284 item(s), 501 transaction(s)] done [0.00s].

sorting and recoding items ... [14 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 3 4 done [0.00s].

writing ... [76 rule(s)] done [0.00s].

creating S4 object ... done [0.00s].

>image(tr)

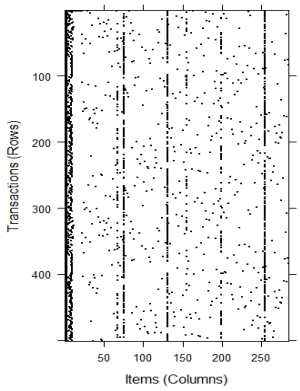


Figure 2.1: Transaction of items

library(arulesViz)

>plot(rules,method="grouped")

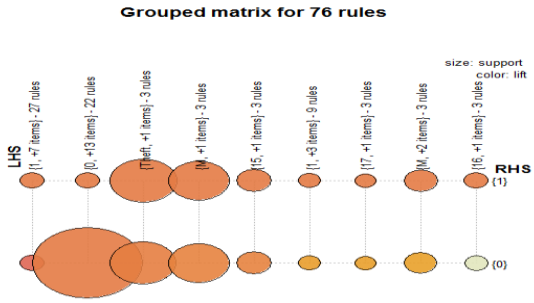


Figure 2.2: Group Matrix

>head(quality(rules));

	support	confidence	lift
1	0.9940120	0.9940120	1.000000
2	1.0000000	1.0000000	1.000000
3	0.1017964	1.0000000	1.006024
4	0.1017964	1.0000000	1.000000
5	0.1177645	0.9833333	0.989257
6	0.1197605	1.0000000	1.000000

plot(rules,measure=c("support","lift"),shading="confidence");

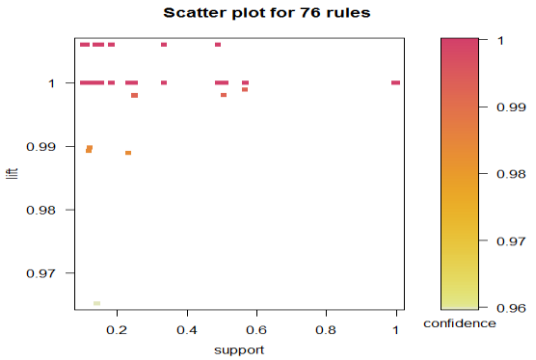


Figure 2.3: Scatter Plot

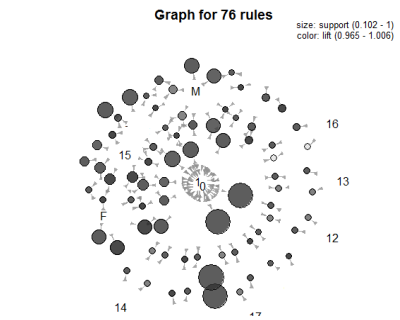
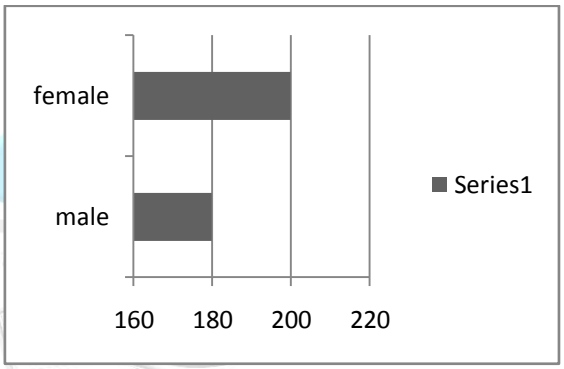


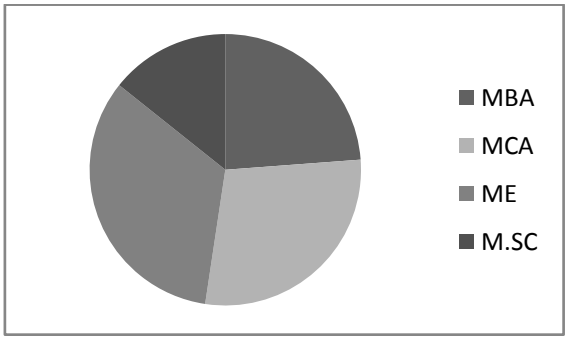
Figure 2.4: Graph for 76 rules

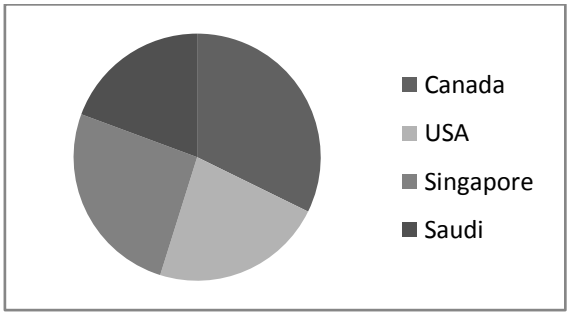
III. FINDING, INTERPRETATION, RECOMMENDATIONS AND SUGGESTIONS

3.1 Finding and interpretations

The gathering of relevant and up-to-date information is a key business process. Information consists of organized facts and figures that have meaning within the context that the information is intended to be interpreted by people. Information is thus a valuable business commodity, and frequently businesses pay money for up-to-date and relevant information.







3.2 Recommendations

- Information collected by the researchers analyzed by using R Tool of Data Mining
- There are three main reasons which make people to move towards foreign countries.
- The reasons are
  - People are moving to improve their status in the society
  - They are in need of more money to fulfill their needs.
  - Most of them were moving for their higher studies.

3.3 Suggestions for further research:

The study encountered several difficulties in reaching a systematic and concrete set of conclusions on the impacts of high skilled emigration from receiving countries. Reliable data are often only available from developed countries, but those sources themselves have limitations. There is little empirical research, either statistical modeling or case studies, that employs comparable methods. First, an objective analysis should be made of the implications of measuring forfeited emigrant skills (brain drain) using different methods or definitions. Second, a systematic review of available data sources should be completed, 53 and a new set of international estimates of loss and selectivity should be constructed. Thirdly, source countries need to improve their data gathering capacity and this may require support and technical assistance from developed countries and international agencies.

REFERENCE

[1]. Chowdhury Mofizur Rahman, Ferdous Ahmed Sohel, Parvez Naushad, Kamruzzaman S. M, "Text Classification Using the Concept of Association Rule of Data Mining," In Proceedings of International Conference on Information Technology, Kathmandu, Nepal, pp Arun K Pujari "Data Mining Techniques", Edition 2001.

[2]. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques" Second Edition. Morgan Kaufmann Publisher, 2006, Pp.123-134.

[3]. Basheer Mohamad Al-Maqaleh and Saleem Khalid Shaab," An Efficient Algorithm for Mining Association Rules using Confident Frequent Itemsets," 2012 Third International Conference on Advanced Computing & Communication Technologies.

[4]. Mihir R. Patel, Dipti P. Rana, and Rupa G. Mehta," FApriori: A Modified Apriori Algorithm Based on Checkpoint,"IEEE 2013.

[5]. Anand H.S. and Vinodchandra S.S.," Applying Correlation Threshold on Apriori Algorithm," 2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN 2013).

[6]. A. Lekha, Dr. C V Srikrishna and Dr. Viji Vinod," Utility of Association Rule Mining: a Case Study using Weka Tool,"2013 IEEE.

[7]. Andreassen, T. W., 1995, "(Dis) Satisfaction with Public Services: the Case of Public Transportation," The Journal of Services Marketing, 9:30-41.

[8]. Arnold, S. J., J. Handelman, and D. J. Tigert, 1996, "Organizational Legitimacy and Store Patronage," Journal of Business Research, 35:229-239.

[9]. Assael, H., 1992, Consumer Behavior and Marketing Action, Pws-Kent Publishing Company, Boston.

[10]. Baker & McKenzie. (2009). Worldwide guide to trade unions and works councils. Retrieved Jan 20, 2012, from http://www.bakermckenzie.com/QRUnderstandGlobalMarketTradeUnionSep09.

[11]. Press trust of India, (2013), "Foreign Investors want progressive, consistent regulations in India: NainaLal Kidwai", at Davos, Switzerland.

[12]. Hind, D. and Moss, S. (2011) Employability Skills. 2nd Edition. Sunderland, Business Education Publishers.