

# Mining Empathic Indicators for Classification of Countries to Alleviate Poverty in the World

Muhammad Shahbaz and Irfan Mushtaq

Department of Computer Science & Engineering, University of Engineering & Technology, G.T. Road, Lahore, Pakistan

Email: Muhammad.Shahbaz@gmail.com, Irfanmushtaq\_mscs@yahoo.com

Aziz Guergachi

Ted Rogers School of Information Technology Management, Ryerson University, Toronto, Ontario, Canada

Email: a2guerga@ryerson.ca

**Abstract**—It is found out that empathy plays a substantially essential role in the economic development and consequently in reducing poverty in the world. In this research our main focus is on human empathy to categorize the nations and compare their empathic level. The aim of this research is to find out the importance of empathic relationships among the countries for the mutual economic development of the countries. For this reason we need countries' statistical data related to economic and social development of the people of those countries, which is collected from the World Bank and United Nations web servers. We also collected data of financial aid given by the countries to other countries for the manifestation of empathic behavior. The empathic indicators are identified and are used for classification of countries using data mining techniques. The unsupervised classification technique "K-mean Clustering" is used to group the data that gave us five clusters of the 118 countries which are compared to one another for their empathic levels. The results indicate some useful and interesting analysis. As a conclusion, it is also suggested that in the future, the same kind of research can be carried out at the organization level.

**Index Terms**—empathy, data mining, clustering, classification, economic indicators

## I. INTRODUCTION

Edith Stein has defined empathy, as a feeling or experience for concern towards others or a foreign consciousness that means a sense of responsibility and positive behaviour that shows someone's inner feelings towards others. Showing empathy or having empathic behavior means to think, feel, and act for the well being of others often without thinking about returns. Being empathic is not just helping others in an hour of need but helping others to cope with their problems without waiting for the foreign hand. Empathy is distinct from sympathy, which can be defined as feeling sorry for others but empathy is feeling the pain of others by placing oneself in the other's situation. Empathy cannot be

labelled as an agreement; rather it is an internal commitment and willingness to look into things from others' eyes [1].

Empathy is the ability to relate to the cognitions, affections, and actions of others and not just being sensitive towards the pain of other's but also to keep others out of those painful situations [2]. Eklund (2011) described that the word empathy has been derived from a Greek word "empathia" that means having insight into the reaction of others i.e. how others will and should react in specific situations. Being empathetic is not limited to having an idea of others reactions in certain situations, rather it is trying to lift others up to the levels where they can themselves become able to overcome their problems [3]. Different theorists have defined empathy differently and some have encompassed more than one concept under the label of empathy. The idea that is shared by all of them is that empathy is a response or perception of a person towards others' experiences [4].

## II. ROLE OF EMPATHY IN ECONOMY

Jeremy Rifkin has described the evolutionary process of communication and energy development along with the psychological and economic development of people [5]. In this book Rifkin has focused on the importance of telecommunication and energy development for bringing people of different parts of the world closer to one another. He says that where the rising levels of communication and increase in energy development are beneficial for human there they also are creating many global problems i.e. climatic alterations, pandemics, and nuclear propagation. Rifkin considers the current scenario of communication and energy regime, predicting that Internet and mobile technology along with small-scale renewable energy commercialization will create an era of distributed capitalism necessary to manage the new energy regime and a heightened global empathy that can help solve global problems [5].

In today's world finance industries are flourishing and income & opportunities in economic field have grown

tremendously. But at the same time poverty is also on an increase in different parts of the world. Most of the manufacturing jobs have been substituted by service sector jobs that are minutely paid. Regressive and non-empathic policies of the past have hindered middle class from progress. There has been no edge of development for working class in the so far economic policies and systems of the world. Most of the economic policies around the world show no empathy for poor. Rich is becoming richer and poor is becoming poorer. This is only because of lack of empathy in world's economic systems. Political systems of most of the countries of the world are standing on wealth and power and there is a scarce realization of the problems faced by the people who are living from hand to mouth. Lack of empathy in world's economic systems has increased the global problem to a greater extent. To overcome these problems we need to build our economic systems and development policies upon some values defined by empathy. On an individual level empathy lets us vaguely appreciate people's feelings but on a group or country level it fosters governments and policy makers to think and secure the wellbeing of others. Our present economic systems do not appreciate the true potential of people and they do not let our leaders to see the true ability of the people. Most of the descriptions of people are not based on the capabilities they possess rather we focus largely on what they are lacking. Rather than appreciating what they can offer, we focus on holes of society and keep on finding out what they need. Our current policies fail to focus on many problems of the people as they are facing them. The results of which is an increase in poverty and problems of people.

Now the question that arises is that how empathy, given a central value in building an economic system, can bring a positive change in the world? The very first benefit that the world will experience will be the changed hierarchy of goals and aims to be achieved. First of all there would be a different set of priorities: raising and educating children, taking care of the elderly, advancing social justice, preserving the planet, and creating effervescent neighborhoods. Work that supports these priorities would be fully valued and rewarded [6].

In this research, the basic aim is to analyze the empathy level of world countries towards other nations. The empathy level is based on the economic development of a country, which shows its empathy according to its economic and development resources. The indicators which are directly related to empathy are described in Table I. From a normal perspective it is observed that the greater the economic stability of a country, the more empathetic the country is; meaning that the country with higher financial development is more empathic than a country with lower financial development.

Being empathic towards others countries has distinct measurements, for example investments of developed countries in poor countries to provide employment to the people and also aiding other countries in disastrous situations.

TABLE I. EMPATHIC INDICATORS [7]

No	Name of Indicator	No	Name of Indicator
1	Human Development Index	15	GNI per capita, Atlas method
2	Food production index	16	Health expenditure per capita
3	IDA grants	17	Imports of goods and services
4	Life expectancy at birth, total	18	Inflation, consumer prices
5	Primary completion rate, total	19	Research and development expenditure
6	Maternal mortality ratio	20	Roads, total network
7	Access to electricity	21	Public spending on education, total
8	Children in employment	22	Unemployment, total
9	Children out of school, primary	23	Malnutrition prevalence, weight for age
10	CO <sub>2</sub> emissions	24	Disaster risk reduction progress score
11	Expenditure per student, primary, Secondary and Tertiary	25	Ease of doing business index
12	Exports of goods and services	26	Electric power consumption
13	GDP per capita	27	Improved water source
14	GINI index	28	Percentage of trained teachers Primary, Secondary and Upper secondary

### III. CLUSTERING

Clustering is one of the central and primitive activity of human beings, which intends to discover natural and hidden patterns in the data. Thus, given a set of data items, cluster analysis aims to provide a framework to put one human being to think and explore insight to the questions like, "can we categorize the data into some groups (clusters)?" If so, then "Into how many clusters we can group the data?" [8].

### IV. DATA PREPERATION

On the World Bank server, in the database of World Development Indicators, the data of 214 World Countries (WC) and territories is available against more than 300 indicators. We first gathered the data of 214 world countries with 40 indicators that were directly related to empathy. We then narrowed down the indicators to ones with the maximum available data of world countries, results in 28 empathic indicators listed in Table I. In order to fill in missing values, smooth out noise and correct inconsistencies we performed the pre-processing mechanism according to the following steps.

Step 1: Ignore the tuple (by setting a threshold of 50%) of WC in which less than 50% of data is available against the empathic indicators.

After step 1, we were left with 118 world countries and 28 empathy indicators.

Step 2: Fill in the missing values by using attributes' averages.

Step 3: Normalize the data by scaling the values between 0.0 and 1.0 using the following equation:

$$v^{\square} = \frac{v - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A$$

where  $v$  is the value of the attribute to be normalized,  $\max_A$  and  $\min_A$  are the maximum and minimum values in one attribute, and  $\text{new\_max}_A = 1$  and  $\text{new\_min}_A = 0$ .

Let empathic indicators ( $EI_n$ ) represent the value of sustainability indicators where  $n=1, 2, 3 \dots 28$ . Let  $WC_n$  represent the world countries where  $n=1, 2, 3 \dots 118$ . Let  $C_1, C_2, C_3, C_4$  and  $C_5$  represent five clusters. Here the number of clusters is set to five after a long trial of experimentation and deciding on the basis of *Davies Bouldin Index* values evaluated by the Cluster Distance Performance Operator of Rapid Minor. For each value of  $EI_n$ , there are 12 values for 12 years (from 2001 to 2012), hence reducing the data from 3D to 2D. These 12 values are collapsed or merged by taking the average value, so for every  $EI_n$  we obtained only one value by combining all the values. Thus, we have 28 different values for every  $WC_n$ . Database view of world countries and indicators is shown in Table II.

TABLE II. DATABASE VIEW OF WORLD COUNTRIES AND INDICATORS

	$EI_1$	$EI_2$	$EI_3$	...	$EI_n$
$WC_1$	$EI_1 WC_1$	$EI_2 WC_1$	$EI_3 WC_1$	....	$EI_n WC_1$
$WC_2$	$EI_1 WC_2$	$EI_2 WC_2$	$EI_3 WC_2$	...	$EI_n WC_2$
$WC_3$	$EI_1 WC_3$	$EI_2 WC_3$	$EI_3 WC_3$	...	$EI_n WC_3$
...	...	...	...	...	...
$WC_n$	$EI_1 WC_n$	$EI_2 WC_n$	$EI_3 WC_n$	...	$EI_n WC_n$

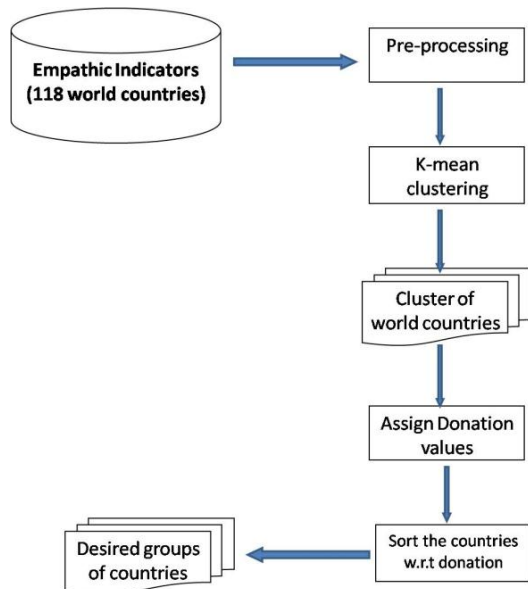


Figure 1. Proposed methodology

$WC_n$  are classified into five clusters for every value of  $EI_n$  by using K-mean clustering method. After deriving the clusters, each  $WC_n$  in each  $C_n$  is assigned the value of the total donation given by that particular  $WC_n$ . Subsequently in every cluster the  $WC_n$  are sorted according to the value of total donations given in

descending order. Lastly, the countries are analyzed and compared for empathic level. The flow chart of the methodology is given in Fig. 1.

## V. RESULTS AND DISCUSSION

### A. Data Manipulation and Processing

The data analysis implementation is done in Rapid Minor 5. 2. A database file consisting of 118 world countries and 28 empathic indicators containing 12 years of data (from 2001 to 2012) is used in this research. The data on indicators is normalized to give equal weightage to all the indicators. The normalized data is then fed into the K-mean clustering algorithm, which produces results in the form of a set of clusters of countries similar to each other based on those 28 indicators.

The final resulting clustered countries by K-means algorithm are shown in the Table III in alphabetical order. The table is truncated to save space. The numbers of countries in different clusters are listed in Table IV.

TABLE III. EMPATHIC INDICATORS

Cluster_1	Cluster_2	Custer_3	Cluster_4	Cluster_5
Afghanistan	Albania	Algeria	Bangladesh	Australia
Angola	Armenia	Argentina	Djibouti	Austria
Benin	Bahrain	Azerbaijan	Haiti	Belgium
Burkina Faso	Barbados	Bhutan	India	Canada
Burundi	Belarus	Bolivia	Kenya	Denmark
Cameroon	Bulgaria	Bosnia	Namibia	Finland
Chad	Chile	Botswana	Nepal	France
Ethiopia	Croatia	Brazil	Nigeria	Germany
Madagascar	Cyprus	China	Pakistan	Iceland
Malawi	Estonia	Colombia	Sudan	Ireland
...	...	...	...	...

TABLE IV. NO OF COUNTRIES IN DIFFERENT CLUSTERS

Cluster No.	No of Countries in that cluster
1	15
2	37
3	29
4	14
5	23

The ultimate goal is not to label these clusters but to compare the empathy level of the countries using a comparison criterion of total donations given by each country. Each country in a group or cluster is at a similar level of development or financial resources.

The donation given by a country shows its empathy in helping the other countries or humans. We also extracted the data of financial aid given by different nations to other countries in various programs or needs. That information is collected from UNCHR's official website for the 13 years (2001-2013) [9].

Next, the countries in all these 5 clusters are arranged in descending order using the aid contributed in 13 years.

TABLE V. CLUSTERS OF WORLD COUNTRIES SORTED WITH RESPECT TO AID

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Angola	Kuwait	Iran, Islamic Rep.	Pakistan	United States
Benin	UAE	Iraq	India	Japan
Mauritania	Saudi Arabia	Morocco	Nigeria	Sweden
Burundi	Korea, Rep.	China	Namibia	Netherlands
Afghanistan	New Zealand	Brazil	Djibouti	United Kingdom
Uganda	Georgia	El Salvador	Yemen, Rep.	Norway
Burkina Faso	Greece	South Africa	Bangladesh	Denmark
Cameroon	Hungary	Algeria	Haiti	Germany
Chad	Portugal	Colombia	Kenya	Canada
Ethiopia	Hong Kong	Argentina	Nepal	Australia
...	...	...	...	...

As it is clear from this information, the country on the very top in each cluster has the same empathy level. Although these countries are donating different amounts to UNHCR, they are donating according to their resources as these are clustered on the basis of those indicators, which are directly related to economic conditions of the countries. In our results, Angola, Kuwait, Iran, Pakistan, and United States are on the same empathy level. As the economic development levels of all these countries are different from one another, they are considered on the same level of empathy because of the ratio of the donation given by them according to their resources.

## VI. CONCLUSION AND FUTURE WORK

In the present research, the clusters of countries obtained after applying K-mean clustering technique were compared on the basis of aid given by those countries to other countries in the time of disaster or whatever they need. Aid is one of the aspects or one of the factors that depicts empathic behavior. However, there are many others.

The investment that countries make in other countries to run that business has an aspect of empathy. These investments and businesses are helpful in raising the quality of living standard of the people in other countries where the investment has been made. This is also a factor or an effort to reduce poverty in those countries by employing people in organizations established in their countries.

Another way of being empathic by the countries to lift poor countries up from the poverty line is letting the people from other countries migrate and settle in the host countries. Through this process, the country being empathic lets the migrants not only lift their own life standards but they can also add to the economy of their own countries by earning and sending money to their lands. This is a rotation of wealth for economic development of other countries. TIER 4 Visa policy

launched by United Kingdom is one of the examples of this kind of empathic behaviour.

Current research was conducted to determine the impact of empathy based behaviour in reducing poverty at the country level. It is suggested that this research work can be used as a basis to conduct future research on an organization level in future using the same strategies which are used in this research. Some of the examples of empathy based behaviour shown by the organizations are as follows.

This methodology can be utilized to study organizations too. Organizations first of all are joined to earn money. The pay scale offered by the organizations for different levels provide better life to the employees, is one of the examples of empathy based behavior at the organizational level. Even a general observation can show us that empathic organizations become more popular as compared to their competitors.

Another instance of empathic behaviour that can raise organizations to the level of success and can lift the life standards of its stakeholders is the concept of 'right man for the right job'. The organizations which let the people do the jobs for which they have no aptitude are not empathic and hence these organizations do not fulfill the demands of the employees and neither of their customers.

## ACKNOWLEDGMENT

The authors would like to thank anonymous referees for helpful comments. This research is partially supported by NSERC

## REFERENCES

- [1] The Collected Works of Edith Stein, *On the Problem of Empathy*, 3rd ed. ICS Publications Washington, D.C. 1989
- [2] W. A. Gentry, T. J. Weber, and G. Sadri, "Empathy in the workplace: A tool for effective leadership," A White paper presented at the Society of Industrial Organizational Psychology Conference, New York, 2007.
- [3] J. Eklund, "Empathy as feeling, understanding and care," Manuscript submitted for publication, Mälardalen University, School of Sustainable Development of Society and Technology, 2011.
- [4] C. R. Rogers, "A theory of therapy, personality and interpersonal relationships, as developed in the client-centered framework," in *Psychology: A Study of Science*, S. Koch, ed. New York: Mc Graw Hill, vol. 3, 1959, pp. 184-256.
- [5] J. Rifkin, "The emptathid civilization: The race to global consciousness in a world in crisis," Jeremy P. Tarcher Inc, 2010.
- [6] E. Cahn, "Empathy, democracy and the economy," Sept. 10, 2013.
- [7] (2013, Aug.). [Online]. Available: <http://www.worldbank.org>
- [8] R. Xu and D. C. Wunsch II, "Recent advances in cluster analysis," *International Journal of Intelligent Computing and Cybernetics*, vol. 1, no. 4, pp. 484-508, 2008.
- [9] (2013, Nov.). [Online]. Available: <http://www.unhcr.org>



**Dr. Muhammad Shahbaz** is working as full professor in the department of Computer Science and Engineering, University of Engineering and Technology, Lahore Pakistan. He completed his masters from Cranfield University, UK in Software Techniques for Computer Aided Engineering and PhD from Loughborough University, UK in Applications of Data Mining in Engineering in 2005. He has authored several

impact factor publications and has attended many international peer reviewed conferences. He is on the reviewer list of several reputed journals and conferences. He has delivered talks and keynote speaks in various universities and conferences. His research interests involved machine learning, data modeling, artificial intelligence, text mining and applications of data mining in health informatics and hydrocarbon exploration. Currently he is leading a research group with his three PhD graduates and several master students working on range of areas including sentiment analysis, players performance optimization in cricket, efficient algorithms for text mining. He is working in collaboration with other research groups around the world including Ryerson University, Salford University, Loughborough University and King Fahad University of Petroleum and Minerals (KUFPM).



**Mr. Irfan Mushtaq** completed his M.Sc from Department of Computer Science & Engineering, University of Engineering and Technology, Lahore, Pakistan. Currently he is a visiting faculty in the Department of Computer Science Foundation University, Rawalpindi. His research interests include data mining and machine learning.



**Aziz Guergachi** is a tenured full professor at Ted Rogers School of Management in Ryerson University. He holds a PhD degree from the University of Ottawa, a Bachelor degree in Engineering from Ecole Supérieure d'Ingénieurs de Marseille, France, and Bachelor of Science in Mathematics from Université d'Aix Marseille I, France. He is a faculty member of Ryerson graduate program in Mechanical and Industrial Engineering, an adjunct professor in the Department of

Mathematics and Statistics at York University and teaches product development and commercialization in Ryerson MBA program