International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 4, No. 1, April 2015, pp. 23~28
ISSN: 2252-8776

Qualitative and Quantitative Modeling of the Artificial Citizens using Genetic Methods

Muhammad Salman, Hajra Klair
Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, KPK, Pakistan

Article Info

Article history:

Received Nov 12, 2014
Revised Feb 20, 2015
Accepted Mar 26, 2015

Keyword:

Artificial population
Forecasting
Genetic algorithm
Genetic simulation
Human development index

ABSTRACT

Human Development Index is a key indicator of a country's development as formulated by the United Nation. It determines a country's progress in three main areas: Life Expectancy, Education and Income. A number of researches have been made for the understanding of how the different factors that together make up the Human development Index are correlated and to what extent each of them affects the HDI value. However, there are always certain unpredictable events that directly or indirectly affect the Human Development Index. Our project aims to incorporate the interdependencies between different factors of the HDI with the simulation of random events in an artificial population to establish the HDI value. This value is future forecasting as the HDI value output would be based upon the input values given by the user, including a certain number of years after which HDI is to be determined. The user would also have other options available to him which he can use to determine the HDI value. That is, after changing the values of various different factors upon which HDI depends by his own choice, the user can effectively study different elements affecting HDI. This can prove to be a great convenience in policy making. Genetic Algorithm is a technique that mimics the process of natural evolution. The essence of Genetic Algorithm is being employed in this system to generate artificial agents for Pakistan's entire population. This population (represented as chromosomes of bit strings) would then undergo various genetic operations like fitness evaluation, cross-over and mutation. Then after certain generations of the initial population have evolved after a specific number of years as input by the user, the Human Development Index value would be calculated. This value would be determined from the statistical analysis of the population. This analysis has been encoded in the system and is result of extensive study of various research articles and books that facilitate in understanding how socio-economic parameters affect each other and the HDI.

Corresponding Author:

Muhammad Salman,
Faculty of Computer Science and Engineering,
GIK Institute of Engineering Sciences and Technology, Topi (23640), Khyber Pakhtunkhwa, Pakistan.
Email: Salman091@gmail.com

1. INTRODUCTION

In a country like Pakistan, it is not to the knowledge of a common person what decisions or events could affect the Human Development Index (HDI) and the standing of a country among the other countries of the world. Dependencies among the economic indicators have been far complex to model resulting in the ambiguity for not only the residents of the society but also for the government officials and major investors in understanding which parameters to address in order to give the country a stable standing in the coming years.

Hence, development of an artificial system is needed, which can study and analyse how certain events and decisions affect the Human Development Index (HDI). The system should be based on real-time data and should be able to produce results accurately.The project targets the Human Development Index (HDI) to predict the effect of policies in the coming years on the masses, simulating the United Nations HDI Model.

The project is challenging not only due to the fact that a large volume of data is considered and used for calculation, which gives rise to several technical and programming issues, but also requires a deep insight on the economic indicators and statistical analysis of the data. The project is designed to serve government organizations as well as private organizations and help them in making vital decisions and policies for a country.

2. LITERATURE REVIEW

The United Nations (UN) calculates the Human Development Index (HDI) based on a yearly survey conducted in all the countries of the world. In order to do so, the UN follows an economic model known as the UN HDI Model. However, except of the three main parameters used in the model to calculate the HDI, which are the Income Index (II), Life Expectancy Index (LEI) and Education Index (EI), the values and the inter-dependencies between the parameters are undisclosed. A lot of research has been done and published on understanding the calculation of the HDI but not much material is available on the inter-dependencies between the parameters which constitute the HDI.

2.1. Definitions and Equations

1. Human Development Index (HDI) is a composite statistic used to measure the quality of life of an individual within a country and rank countries by level of "human development".

$$\mathrm{HDI} = \sqrt[3]{\mathrm{LEI} \cdot \mathrm{EI} \cdot \mathrm{II}}.$$

2. Life expectancy is the expected (in the statistical sense) number of years of life remaining at a given age. In HDI, Life expectancy at birth is taken into consideration which is calculated from a life table (which comprises values of birth rate and death rate of the total Population).

   Life Expectancy Index (LEI) = (LE – 25) / (85 - 25)

3. Education Index is measured by the adult literacy rate (with two-thirds weighting) and the combined primary, secondary, and tertiary gross enrolment ratio (with one-third weighting).

   Education Index (EI) = (2/3 X ALI) + (1/3 X GEI)

4. The Adult literacy index (ALI) is a statistical measure used to determine how many adults can read and write in a certain area or nation.

   Adult Literacy Index (ALI) = (ALR – 0) / (100 – 0)

5. The Gross Enrolment Ratio (GER) or Gross Enrolment Index (GEI) is a statistical measure used to determine the number of students enrolled in school at several different grade levels (like elementary, middle school and high school), and examine it to analyse the ratio of the number of students who live in that country to those who qualify for the particular grade level.

   Gross Enrolment Index (GEI) = (CGER – 0) / 100 – 0

6. Gross domestic product (GDP) refers to the market value of all final goods and services produced within a country in a given period.

   Income Index (II) = (log (GDPpc) – log(100)) / (log(4000) – log(100))

7. PC stands for per capita which means per person.
8.

2.2. Model

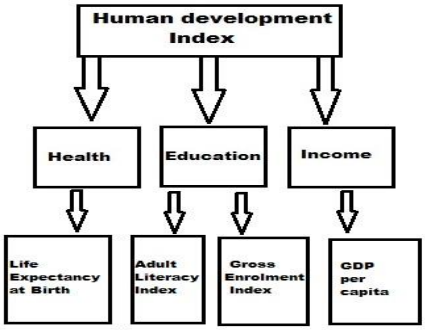

Figure 1. Model for calculating HDI Value

2.3. Adjustment Factor of Socio-Economic Factors

Beside the three main parameter of HDI, we are also giving user the flexibility to enter ten most important socio-economic factors which directly or indirectly affect HDI. Those factors include suicide attacks, floods, earthquakes, warfare, storms/cyclones, cancer, HIV/AIDS, hepatitis, malaria, and road accidents. As a result, there is an increase in death rate and the income of individuals is also affected negatively. An adjustment factor is calculated based on these socio economic factors, which is later multiplied with the HDI value before the result is displayed to user.

We have formulated this affect in our own way based on the percent increase in death rates plus percent decrease in income (GNI per capita).

3. METHODOLOGY

A Genetic Algorithm follows the biological process of Genetics, giving birth to a population of chromosomes and performing biological functions such as crossover and mutation. The chromosomes are then evaluated on the basis of a fitness function, eliminating the ones not passing the fitness test, hence, leaving behind a set of chromosomes which constitute one of the probable and possible solutions of the problem. In this project:

1. The chromosomes represent the human beings in a country. Hence, for Pakistan, initially 180 Million chromosomes are generated based on the current statistics of Pakistan.
2. Each chromosome is an agent, with its own attributes, and all of these chromosomes form a 'population'. Each chromosome consists of seven bits of string in form of 1010101 as shown below in Figure 2.
3. After the generation of chromosomes, we now apply values of all factors on which HDI is dependent upon i-e Life Expectancy Index, Education Index, and Income Index.
4. Further, ten socio-economic factors which are mentioned above can be entered in the software now.
5. After all these factors are fed, by inserting number of years, as per required, the user can run the country for the desired number of years. For example, one year entered is equivalent to one generation.
6. Now, the biological process of genetics are applied on each chromosomes based on birth rate, death rate, crossover rate, and mutation rate.
7. Each chromosome is evaluated on the basis of the fitness criteria. This fitness criterion is based on how and to what extent an agent can be affected by a particular operation (which is a simulation of a particular policy or a real life event).
8. This procedure is continued till the number of years specified is reached.

After the number of generations specified is reached, the software is expected to have produced a set of chromosomes constituting the predicted population, economic reforms, educational plans, effective resource utilization, eradication of poverty, decline in growth rate, terrorism, unforeseen events like floods, earth quakes, etc. with great accuracy, provided the inter-dependencies between the parameters are well laid out.

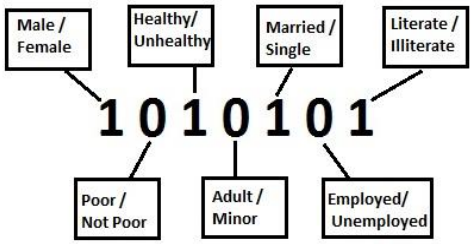

Figure 2

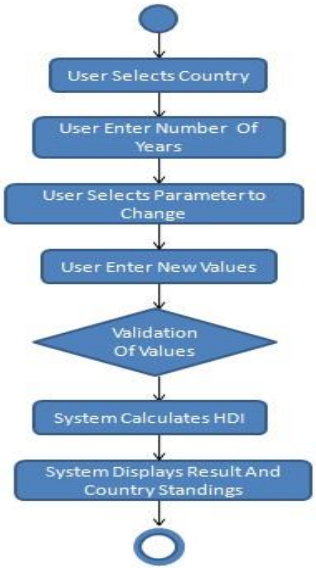

Figure 3. Flow of Software in Order of its Operation is Shown Below

4. RESULTS

The program was run for the data set as jotted down in the previous section. The results show that the HDI gradually increases along with the three main factors that constitute the HDI. The value of the HDI increases by a factor of 12.65% in the coming ten years. The results calculated are based on the inter-dependencies and values studied and gathered from various sources, and include random and unknown factors such as suicide attacks, floods, road accidents etc. The next section describes the flexibility of the program to adapt different case scenarios. Graphical Result produced by the software after the specified number of years is shown below in Figure 4:

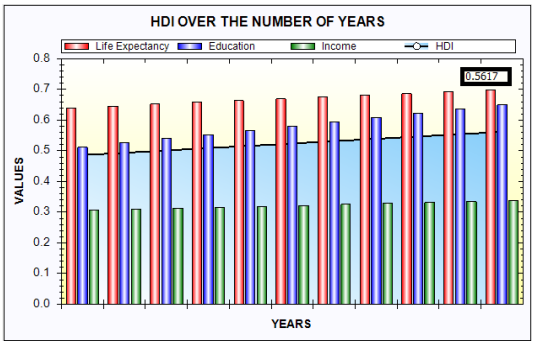

Figure 4. Graphical Result produced by the software after the specified number of years

5. CONCLUSION

The program provides the user with the flexibility and user-friendliness to change the case scenarios in order to achieve the outcomes of desired interest. The program allows the users to select their country of interest which makes the project limitless on one country's grounds. The program also allows the users to enter and/or modify the different parameters according to their desired interests. The program also targets the unknown and random factors such as the suicide attacks, floods, road accidents, various diseases etc, and lets the user to either set yearly rates or set manual values as per his desired interests. Hence, the program constitutes results based on the user input and provides users with a user-friendly Graphical User Interface.

Considering the unlimited number of parameters and their unlimited inter-dependencies, it becomes an impossible task to predict the results with a 100% accuracy rate. However, the accuracy can be increased by increasing more parameters and studying their inter-dependencies with other parameters. Hence, a lot of dedicated research work is required. The project does not only demand the knowledge of statistical analysis and economic indicators, but also requires several programming techniques since the calculation needs to be performed on a very large data. Also, for countries with larger populations, the technique of Cloud Computing is recommended as it serves as the best solution for faster result generation and calculation

ACKNOWLEDGEMENTS

I wish to express my deep sense of gratitude to our Project Advisors, Dr. Zeeshan-ul-Hassan Usmani, Project Head GSDM, Interactive Group, and Dr. Taimur Qureshi, Project Manager, Interactive Group for their able guidance and useful suggestions, which helped us in completing the project work, in time.

Needless to mention that Faculty of Computer Science, GIK Institute of Engineering Sciences and Technology, Topi , Khyber Pakhtunkhwa who had been a source of inspiration and for his timely guidance in the conduct of our project work. Finally, yet importantly, I would like to express my heartfelt thanks to our respective beloved parents for their blessings, our friends/classmates for their help and wishes for the successful completion of this project.

REFERENCES

[1] National Institute of population Studies Islamabad Pakistan demograhpic and health survey 2006-07, Prof. Dr. Muhammad Aslam, "Muhammad Aslam Chaudry Revamping of vital Events System in Pakistan", Department of Statistics Quaid-I-Azam University, October, 2008
[2] Government of Pakistan, "Statistics Division and Federal Bureau of Statistics Pakistan Labour force survey 2009-2010 (Quaterly reports)"
[3] United Nations Human Development Report 2010 20th Anniversary Edition
[4] Ramirez, A., Ranis, G., Stewart, F, "Economic Growth and Human Development Yale University Economic Growth Centre, QEH", *Working Paper Series*, No. 18, 1998
[5] Ranis, G. and Stewart, F, "Growth and Human Development: Pakistan in Comparative Perspective The Pakistan Development Review", (*Winter*), 36:4, Pp.333-352, 1997
[6] Life expectancy : "Information on Public Health Observatory recommended methods"
[7] Chiang CL., "The life table and its applications*"*, Malabar (FL): Krieger; 1984.
[8] Jana Asher, Beth Osborne Daponte, "A Hypothetical Cohort Model of Human Development"
[9] Nhat M, "Nguyen Purchasing Power Parity IFM", Section 01 Fall 2005
[10] Wei R, Curtin LR, Arias E, and Anderson RN. U.S. "Decennial Life Tables for 1999-2001; Methodology of the United States Life Tables", *National Vital Statistics Reports*, 2008, 57:4.
[11] Calculating life expectancy and infant mortality rates Mapping health inequalities across London: technical supplement
[12] Ramirez, A., Ranis, G., Stewart, F., "Economic Growth and Human Development", Yale University Economic Growth Centre, QEH, *Working Paper Series*, No. 18, 1998
[13] Ranis, G. and Stewart, F. (1997) Growth and Human Development: Pakistan in Comparative Perspective, The Pakistan Development Review, (*Winter*), 36:4, Pp.333-352
[14] http://www.erpho.org.uk/viewData.asp?itemID=5951
[15] http://www.gad.gov.uk/Life_Tables/Interim_Life_Tables.htm
[16] http://www.unicef.org/infobycountry/stats_popup1.html
[17] en.wikipedia.org/wiki Human_Development_Index.htm
[18] http://www.nationsonline.org/oneworld/human_development.htm
[19] http://www.nationsonline.org/oneworld/PPPGNI.htm
[20] en.wikipedia.org/wiki/Purchasing_power_parity.htm

[21] en.wikipedia.org/wiki/GNI.htm
[22] en.wikipedia.org/wiki/Life_expectancy.htm
[23] http://internationalecon.com/Finance/Fch30/F30-1.php
[24] Health Canada 2005. First Nations comparable health indicators. <www.hc-sc.gc.ca/fnih-spni/pubs/gen/2005-01_health-sante_indicat/index_e.html>.
[25] www.findem.com.au/lifeexpectancy.pdf
[26] http://moneycentral.msn.com/investor/calcs/n_expect/main.asp
[27] http://www.who.int/whosis/whostat/en/
[28] The Measure Of America by sarah Burd-Sharps, Kristen Lewis, And Eduardo Borges Martins
[29] The Process Of Economic Development by James M. Cypher, James L. Dietz

BIOGRAPHIES OF AUTHORS

Muhammad Salman is currently working as a Software Engineer in iResearch Department, Interactive Group, Islamabad, Pakistan. He received his BS degree in computer software engineering from GIK Institute of Engineering Sciences and Technology, Topi (23640), Khyber Pakhtunkhwa, Pakistan. His research interests include Artifical Intelligence, Data Mining, cloud computing and distributed computing.

Hajra Klair is currently working as a IT Marketing Specialst in Elixir Technologies, Islamabad, Pakistan. She received her BS degree in computer software engineering from GIK Institute of Engineering Sciences and Technology, Topi (23640), Khyber Pakhtunkhwa, Pakistan.