

# Comparative studies on Sustainable Development Goals(SDG) in India using Data Mining approach

P. Rajesh<sup>1\*</sup> and B. Santhosh Kumar<sup>2</sup>

<https://doi.org/10.56343/STET.116.014.002.006>

<http://stetjournals.com>

## Article History

Received: 18-08-2020

Revised and Accepted : 14-10-2020

Published: 28-12-2020

## Abstract

The sustainable development goals (SDGs) are a universal plan for unique assessment for all countries to find the rankings likes, end of poverty, protecting the planet, and ensuring prosperity for all. Now a days the research area of SDGs is one of the growing fields in worldwide research, particularly in the field of forecasting the future SDGs ranks. The main objective of this paper is analyzing the SDGs by various independent metrics in the states of Tamil Nadu, Kerala and Karnataka in India and by taking consideration into three different state SDGs index using data mining and statistical approaches for retrieving various hidden information. Numerical illustrations are also used to prove the proposed results.

**Key words:** big data applications, data mining, forecasting, poverty eradication, SDGs

## INTRODUCTION

Data mining is the process of analyzing hidden patterns for using pre-existing data. Data mining is also known as data discovery and knowledge discovery for handing advanced data analysis (Jiawei., *et al.* 2011). The major steps involved in a data mining process namely locate the data, data collection, data cleaning, integration, data selection, data transformation and discovering the knowledge (Bocca., *et al.* 2016)

Big data applications may offer the ability to collect and analyze 'real time' information from across Environment and Development Policy Section (ESCAP)'s 62 member States for policies that relate to the 2030 Agenda's 17 goals and their 169 targets (Maarroof, 2015). Agenda 2030 was presented at the

Senior Professional Forum for Strategic Planning (Elkin and Katz, 2019).The United Nations' 'No Poverty' goals seeks to eradicate poverty in all its forms through the achievement of seven key targets, each monitored by various indicators, making it imperative for countries to report poverty data that can be compared over time, and are available at relatively frequent intervals (Lokanathan *et al.*, 2017). This study investigates, at the country level, the adverse effects of changes in metal inputs on the achievement of Sustainable Development Goals (SDGs) in three Indian states. It also highlights various socio-economic consequences that urgently require decoupling in order to achieve the SDGs (Nansi *et al.*, 2019).

In data mining techniques, normalization is one of the most important concepts for prepare a well suitable dataset with unique format. After using the normalization techniques various scales of information converted into similar scale of information. Various normalization techniques are also used to handling the data analysis, one of the most popular normalization techniques called maxima and minima normalization (Teixeira and Stephany, 2013; Rathod

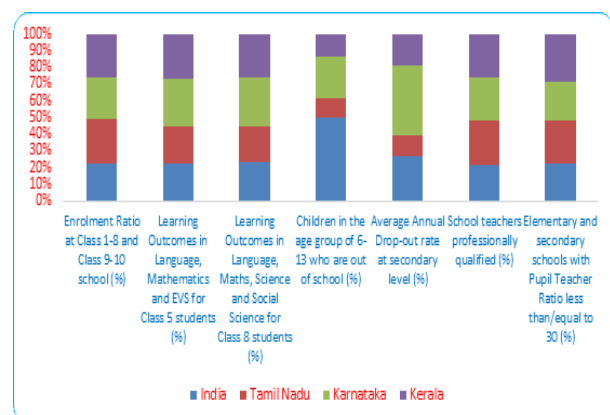


**P. Rajesh**

email: [rajeshdatamining@gmail.com](mailto:rajeshdatamining@gmail.com)

<sup>1</sup>PG Department of Computer Science, Government Arts College, C. Mutlur, Chidambaram, Tamil Nadu, India.

<sup>2</sup> Department of Computer Applications, Periyar Arts College, Cuddalore, Tamil Nadu, India.



**Fig. 1.** Comparison of Quality Education dataset (SDG - 4) in three different states with that of India

*et al.*, 2018; Rajesh and Karthikeyan, 2018; 2019a,b and c; Rajesh *et al.*, 2019)

## MATERIALS AND METHODS

Analysis of No Poverty dataset (SDG1) include the parameters namely, Population living below National Poverty line (%), Household's members covered namely health insurance (%), Demanded employment under MGNREGA (%), Social protection benefits under

Maternity Benefits, Number of homeless households per 10,000 households. Analysis of Zero Hunger (SDG2) include the following parameters, namely, Rural households' monthly income of less than Rs.5,000, Children under age 5 years who are stunted (%), Pregnant women aged 15-49 years who are anaemic (11.0g/dl) (%), and Rice, wheat and coarse cereals produced annually per unit area (Kg/Ha).

**Table 1:** Comparison of No Poverty dataset (SDG - 1), in three different states with that of India

State	Population living below National Poverty line (%)	Households' members covered under health insurance (%)	Demanded employment under MGNREGA (%)	Social protection under Maternity benefits	Number of homeless households per 10,000 households	Mean (SDG-1)
India	21.92	28.7	84.75	36.4	10.39	36.432
Tamil Nadu	11.28	64.1	98.83	29.5	4.56	41.654
Karnataka	20.91	28.1	84.26	19.9	8.45	32.324
Kerala	7.05	47.7	87.98	20.4	3.4	33.306

**Table 2:** Comparison of Zero Hunger dataset (SDG - 2), of three different states with that of India

States	Rural households' monthly income with less than Rs.5,000	Children under age of 5 years who are stunted (%)	Pregnant women aged 15-49 years who are anaemic (11.0g/dl) (%)	Rice, wheat and coarse cereals produced annually per unit area (Kg/Ha)
India	1.01	38.4	50.3	2509.22
Tamil Nadu	0.8	27.1	44.4	3788.49
Karnataka	1.1	36.2	45.4	2157.83
Kerala	0.74	19.7	22.6	2788.79

**Table 3:** Comparison of Good Health and Well-being dataset (SDG - 3), of three different states with that of India

States	Maternal Mortality Ratio	Under-Age five mortality rate per 1,000 live births	Children aged 12-23 months fully immunized (BCG, and other vaccines) (%)	Annual notification of Tuberculosis cases per 1 lakh population	Number of governmental physicians, nurses and midwives per 1,00,000 population
India	130	50	62	138.33	220.96
Tamil Nadu	66	27	69.7	119	426.04
Karnataka	108	32	62.6	123	452.93
Kerala	46	7	82.1	67	762.13

Analysis of Good Health and Well-being(SDG3) include the following parameters namely, Maternal Mortality Ratio,Under-Age five mortality rate per 1,000 live births,Children aged 12-23 months fully immunized (BCG, and vaccine) (%),Annual notification of Tuberculosis cases per 1 lakh population,Number of governmental physicians, nurses and midwives per 1,00,000 population.

Analysis of Analysis of Zero Hunger (SDG4) include the following parameters, namely, Enrolment Ratio at Class 1-8 and Class 9-10 school (%),Learning Outcomes in Language, Mathematics and EVS for Class 5 students (%),Learning Outcomes in Language, Maths, Science and Social Science for Class 8 students (%),Children in the age group of 6-13 who are out of school (%),Average Annual Drop-out rate at secondary level (%),School teachers profession qualified (%) and Elementary and secondary schools with Pupil Teacher Ratio less than/equal to 30 (%).

## RESULTS AND DISCUSSION

The Sustainable Development Goals (SDGs) called as Global Goals are the collection of 17 built in interlinked global goals developed by the United Nations General Assembly in 2015 and are initiated to be achieved by the year 2030. Using SGDs, we ranked three states within India using 17 goals. Table 1 indicates the SDG-1 namely “no poverty” occurring in India to be only 36.43%. At the same time the Government of Tamil Nadu has higher value of 41.65% compared to Karnataka and Kerala, and India as a whole.

Analysis of Zero Hunger dataset (SDG - 2) showed that Tamil Nadu has higher values for the production of rice, wheat and coarse cereals while the state of Karnataka has higher values in the other three parameters (Table 2).

Analysis of Good Health and Well-being dataset (SDG - 3) showed that Kerala has better achievements in annual notifications of Tuberculosis cases per 1 lakh population, while Karnataka has better results in the other three parameters (Table 3).

Analysis of Quality Education dataset indicate Kerala has the highest percentage in the category of Education compared to Tamil Nadu and Karnataka. In this case the school entry enrollment 92% compared to other states.(Fig. 1)

## REFERENCES

Bocca, F. F., Henrique, L.and Rodrigues,A., 2016. The effect of tuning, feature engineering and feature selection in data mining applied to rainfed sugarcane yield modelling. *Comput. Electron. Agric.*, 128: 67-76.

- Elkin, Z. and Katz, I., 2019. Implementation of the Sustainable Development Goals: National Review. *Glob. Risk Insights*, 430.
- Jiawei Han, Micheline Kamber and Jian Pei, 2011. *Data Mining: Concepts and Techniques*. The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers., US.
- Lokanathan, S., Perera-Gomez, T., and Zuhyle, S., 2017. Mapping Big Data Solutions for Sustainable Development Goals. *LIRNEasia, Sri Lanka*. 10: 1-81.
- Maarroof, A., 2015. Big Data and the 2030 Agenda for Sustainable Development. *Unescap*, 1-55.
- Nansai, K., Kondo, Y., Giurco, D., Sussman, D., Nakajima, K., Kagawa, S., Takayanagi, W., Shigetomi, Y., and Tohno, S., 2019. Nexus between economy-wide metal inputs and the deterioration of sustainable development goals. *Resour. Conserv. Recycl.*, 149: 12-19.
- Rajesh, P., 2020. Developing Data Mining Based Stochastic Model for Investigate the Parameters for Increasing Wheat and Groundnut Productions. *Adalya. J.*, 9(2): 319-327.
- Rajesh, P. and Karthikeyan, M., 2018. Predication of Labour Demand in Agriculture Based on Comparative Study of Different Data Using Data Mining and Stochastic Approach. *Int. j. eng. sci. invention*, 2: 86-97.
- Rajesh, P. and Karthikeyan, M., 2019a. Data Mining Approaches to Predict the Factors that Affect the Agriculture Growth using Stochastic Model.*Int. j. comput. sci. eng.* 7(4): 18-23.
- Rajesh, P. and Karthikeyan, M., 2019b. Data Mining Approaches to Predict the Factors that Affect the Groundwater Level using Stochastic Model. *AIP Conference Proceedings*. 2177: 020079-1-020079-11.
- Rajesh, P. and Karthikeyan, M., 2019c. Data Assimilation of Gross Domestic Product (GDP) in India using Stochastic Data Mining Approach. *J. Comput. Theor. Nanos.*, 16(4): 1478-1484.
- Rajesh, P. and Karthikeyan, M. and R. Arulpavai, R., 2019. Data Mining Algorithm to predict the Factors for Agricultural Development using Stochastic Model. *IJRTE*, 8(3): 2713-2717.
- Rathod, S., Singh, K. N., Patil, S. G., Naik, R. H., Ray, M. and Meena, V. S. 2018. Modeling and forecasting of oilseed production of India through artificial intelligence techniques.*Indian J. Agric. Sci.* 88(1): 22-27.
- Serrano,J. M., Delgado, G., Aranda, V., Calero, J. and Vila, M. A., 2009. Using fuzzy data mining to evaluate survey data from olive grove cultivation.*Comp. Elect. Agr.*,65(1): 99-113.
- Teixeira de Lima, G. R. and Stephany,S. 2013. A new classification approach for detecting severe weather patterns. *Comput. Geosci.*, 57: 158-165.