A Clustering Based Performance Evaluation System for MGNREG Scheme

Ch. Geethanjali[1], V. Valli Kumari[2]

Dept. of Computer Science and Systems Engineering, Andhra University, Visakhapatnam, Andhra Pradesh[1]

Professor, Dept. of Computer Science and Systems Engineering, Andhra University, Visakhapatnam, Andhra Pradesh[2]

Abstract: The Indian government is introducing many schemes for the welfare of poor but they are not reaching them in a proper and effective way. There is an uncertainty or lack of performance analysis in implementation of schemes. Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) is one of such schemes implemented by the government of India which aims at providing employment to the poor in rural areas by assigning local works to them. Though the scheme is implemented throughout the country, there is uncertainty in analyzing its impact on rural people. So a performance evaluation system is useful for the government to analyze performance and success of scheme in villages. This paper describes a performance evaluation system on the data collected from www.nrega.gov.in web portal for the district of Visakhapatnam. Suitable statistical and data mining techniques are applied on this data to evaluate the performance of scheme for the district of Visakhapatnam.

Keywords: Clustering, Enhanced k-means, MGNREGS.

I. INTRODUCTION

In India majority of the population are living in rural areas. The government of India has announced many schemes for rural masses. The main aim of these schemes is to bring them out of the folder of poverty. Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) is a pioneer scheme for providing minimum 100 days of job [1] to the rural and tribal population, whose livelihood depends on the daily wages. Family members living in the same village whose age is above eighteen need to enrol in the job card, to become a registered member of the scheme. Each of the micro family is given an unique job card. After the registration they will get work within 15 days, else unemployment allowance has to be provided.

The core objective of this scheme is to enhance livelihood and economic security for the people of rural areas by providing at least one hundred days of guaranteed wage employment on demand. This work guarantee is also serving other objectives [4]: generating productive assets, environment protection, rural empowerment of women, reducing migration from villages and strengthening social equity, development initiative and public investments for creation of long lasting assets with a decentralized implementation. This scheme is expected to improve people's livelihood by creating and developing economic and social infrastructure in rural areas.

The works taken up under this scheme
1. Water conservation and water harvesting.
2. Drought proofing
3. Irrigation of canals including micro and minor irrigation work.
4. Renovation of traditional water bodies.
5. Flood control and protection works.
6. Rural connectivity to provide all-weather access.
7. Any other work which may be notified by the government.

Here a performance evaluation system is useful for local, state and central governments to analyze the impact of the scheme in the villages. For this Enhanced k-means clustering technique and linear regression technique is applied on the data.

Clustering is one of the primary data analysis methods which is also known as unsupervised learning. It is the process of organizing data objects into a set of disjoint classes called clusters. This Cluster analysis partitions a given data set into groups based on specified features so that the data points within a group are more similar to each other than the points in different groups. K-means is one of the widely used algorithms which classifies a given dataset through a certain number of clusters (k clusters).

II. RELATED WORK

Abhishek Thakur [2] the objective of this study is to show the change in relationship between labourers and farmers by the implementation of MGNREGS. This study is conducted in Seoni District of Madhya Pradesh to understand the source of livelihood and impact of the scheme.

Basu Arnab K [3] this paper provides a theoretical framework for the evaluation of a number of conflicting observations and empirical results on the impact of an employment guarantee scheme on agricultural wages and employment.

Jyoti poonia [5] this paper shows India's approach to social protection since independence. It looks at how gender concerns have been addressed within public works, and specifically in the MGNREGS guidelines.

S.Krishnan [7] analyses this is the most successful scheme to eradicate poverty in India. It provides money directly in the hands of poor, especially women without any middlemen.

III. PERFORMANCE EVALUATION SYSTEM

In this proposed system an attempt has been made to develop a performance evaluation system to analyze the impact of the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) in different villages. Here the effectiveness of scheme is analyzed in two ways, the first one is clustering the villages using Enhanced k-means algorithm and the next one involves predicting next year value by considering previous year data.

A. Prediction of Households Working/Average Wage/Budget

Based on the past history i.e. from 2008 to 2014 years, the number of households who worked in their respective mandals may work in the future i.e. in year 2015 – the trend is calculated for each mandal. Expected number of households who may work in the future (trend) years i.e. 2015 is showed along with previous year data. And average wage rates and budget are also predicted in respective mandals for the year 2015 by considering their previous year data. Here linear regression technique is applied on the mandal data for prediction. This may provide guidelines to government that how much to spend on what mandal.
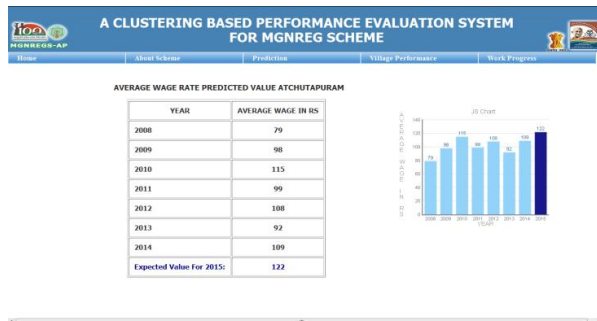

Fig 1 Prediction output

Fig 1 depicts trend changes from 2008 to 2015 for atchutapuram mandal of Visakhapatnam district. It also provides yearwise graphical representation of households working/average wage/budget along with predicted value.

B. Evaluating Village Performance by Applying Enhanced k-means Algorithm

For evaluating village performance mandal data of Visakhapatnam district is collected from nrega [6] web portal. This web portal is developed by Department of Rural Development, Government of Andhra Pradesh which provides data regarding MGNREG scheme. Clustering technique Enhanced k-means is applied on the collected data. This algorithm clusters the villages into best, average and below average for the selected mandal by considering the parameters of village like number of households working, their average wage rates, work progress rate etc.

Enhanced k-means clustering algorithm

Input

D = {d1, d2,......,dn} // set of n data items.
k // Number of desired clusters.

Output

A set of k clusters.

Steps

1. Identify maximum and minimum element in each column of dataset and determine range as the difference between these elements;
2. Find the column having the maximum range;
3. Sort the entire data set in increasing order based on the column having the maximum range;
4. Partition the sorted data set into 'k' equal parts;
5. Determine the arithmetic mean of each part obtained in Step 4 as c1, c2,....ck;
These mean values will be considered initial centroids of data set.

6. Repeat

   6.1 Assign each data item di to the cluster which has the closest centroid;

   6.2 Calculate new mean of each cluster;

Until convergence criterion is met.

A data point is assigned to its closest centroid by calculating Euclidean distance. The Euclidean distance between two data points X = (x1, x2, x3... xm) and Y = (y1, y2, y3... ym) can be calculated using the following formula:

$$D(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_m - y_m)^2}$$

Convergence is a situation where the centroids of clusters do not change.
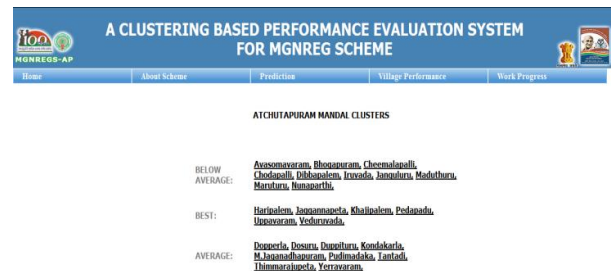

Fig 2 Clustering output

Fig 2 depicts villages of selected mandal, and they are grouped into clusters as best, average and below average. For this Enhanced k-means algorithm is applied on mandal data.

In the original k-means algorithm initial centroids are selected randomly, where as the enhanced algorithm calculates initial centroids in a more meaningful way, in accordance with the distribution of data. So, the algorithm converges faster than the original kmeans algorithm. Since the method for calculating the initial centroids is based on the technique of sorting the data, it requires less time compared to original k-means. The sorting technique used in this algorithm is heap sort which takes O (n logn) to sort the data.

C. Work Progress Study

This study is used to analyze the available data in the database and generate reports for monitoring the successful implementation of the scheme. For this

yearwise data of work progress and work category analysis is showed using graphical representation.
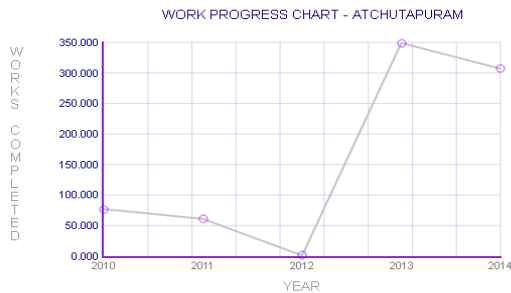

Fig 3 Yearwise works completed analysis

Fig 3 is plotted by taking year value on x-axis and number of works completed value on y-axis. The above graph shows that number of works completed is maximum in the year 2013 and minimum in the year 2012.
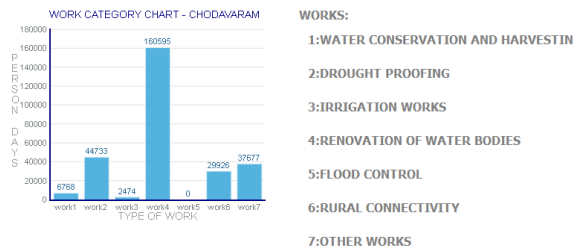

Fig 4 Person days generated for different categories of works

Fig 4 depicts person days generated for different types of work for the given mandal. The above graph is plotted by taking type of works on x-axis and number of person days generated on y-axis. From the graph it is observed that the person days generated for renovation of water bodies work is more where as zero for flood control work.

IV. CONCLUSION AND FUTURE ENHANCEMENTS

MGNREGS is the major socio-economic reform scheme implemented throughout the country. By collecting various factors which influence in implementation part, we have studied the effectiveness of the assets created and their contribution towards long term development and sustained employment generation in the villages as well as issues of wages, processes of implementation, etc. For this linear regression and Enhanced k-means clustering techniques are applied on the Visakhapatnam district data and analyzed the performance.

The proposed method can be enhanced by considering more number of parameters in the implementation of scheme and assessment of performance can be extended to various domains like agriculture, impact on changes of forest resources, ensuring better food security to beneficiaries, examining the pattern of migration from rural to urban, etc

REFERENCES

[1] Lakshmi Vaidyanathan, Evaluation of National Rural Employment Guarantee Act (NREGA) in Tamilnadu, Indian Institute of Technology, Madras, India, 2009.
[2] Abhishek Thakur, Dr. Manish K. Jha (2011), A Study on MGNREGA and its impact on wage and work relation.
[3] Basu Arnab K., Impact of rural employment guarantee schemes on seasonal labour markets: optimum compensation and workers' welfare, College of William and Mary, mimeo, 2007.
[4] Soumya Mohanty, Dr. Nihar Ranjan Mishra (2012), Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) and Tribal Livelihoods: A Case Study in Sundargarh District of Odisha.
[5] Jyoti poonia (2012), "Critical Study of MGNREGA: Impact and women's participation," International Journal of Human Development and Management Sciences, Vol. 1 No. 1.
[6] Information about MGNREGS from Andhra Pradesh rural development website, www.nrega.ap.gov.in.
[7] S.Krishnan, Dr. A.Balakrishnan (2014), "MGNREGA Marching towards achieving the millennium development goals-an analysis," Journal of International Academic Research for Multidisciplinary. Volume 2, Issue 1.