WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# AN OVERVIEW TO PREDICT POVERTY FROM SATELLITE IMAGES USING MACHINE LEARNING

***Vinay Raj A S[1], Ajay S[2], Bhanu Prasad Yadav P[3], Charan Prajwal[4], Chethan M A[5]***

**[1]Assistant Professor, Department of Computer Science Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.**
**[2,3,4,5]Undergraduate Scholar, Department of Computer Science and Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.**

A B S T R A C T

Poverty level prediction is determined as one of the major complex and crucial aspects in the world, the main esteem of the project is to estimate the level of poverty by the help of satellite images one of the major issues is predicting the poverty level manually by visiting every rural area and it is considered to be time consuming and labour involving process. Some of the complex algorithms such as recurrent neural networks and convolutional neural networks will help in determining the amount of poverty level based on the training we provide and predicting by considering various parameters such as the type of roof, industrial area, rural area and also based on the amount of water supply as well as agricultural production. The machine uses various datasets which will help it in understanding and training itself by the help of different types of learning such as reinforcement, supervised and unsupervised learning. Hence there is an alternative approach to solve the problem by reducing a massive amount of work and not making the government to much struggle in figuring out the problem for predicting the poverty level

Keywords: Regression, Poverty Prediction, Feature Selection, Recurrent Neural Network, Convolutional Neural Network, Deep Learning.

## 1. Introduction

One of the main goals of the development goals sustainable 17 is to tackle the problem of poverty and this has been very exhaustive process in determining the solution for this problem, many researchers and government authorities are trying in different ways to solve and tackle the problems. These are Considered to be important complex, social and political issues. There are many lacks of effective approaches in solving them, there are also new approaches in solving them such as the source data, imagery data, street view and other forms of information data and this way of learning uses many types of languages like machine learning, deep learning and some other forms of analysis of poverty and these types of outputs may differ during night time and day time as well the output may vary for both due to certain different factors like daylight and other forms.
In the initial stages artificial intelligence was considered to be one of the mainly used tools for any kind of problems but due to the recent developments many different tools have been built up by which the advantages in solving any problem increases rapidly so there are many different ways in solving these issues, this is one of the common and usual approaches followed and by the help of algorithms of deep learning which help in determining complex type of problems when compared with machine learning.

## 2. Literature Review

### *2.1 Multidimensional poverty and counting poverty measurement*

Firstly, the methods involved in measurements are more largely dependent and the variables are cardinal in fact there are many different dimensional of interest or categorial and the second methods are still under study and Multidimensional poverty and counting poverty measurement has excited many researchers and has given attention to the research and development organizers. There are many different unidimensional Space but the recent tasks and the recent efforts that are put to identify and brought to shown that Multidimensional poverty and counting poverty measurement are working powerful and has brought impact.
The main goal of this reference paper which was researched is to bring a new and better way methodologies that have been found in this paper so the counting and multidimensional approach are in a better way used to reflect and deeply find out the multidimensional poverty level and this has brought a dept knowledge of severity in finding the solution.
Some of the measures used in this satisfy in a range and make use of many different properties. One of the important points to be considered are developing these kinds of new methodologies with the help of new data, by the help of this reference paper it was shown that this methodology has proposed many useful approaches and this can be applied to any real data. Some of the basic things to noted are the multidimensional property has very good headcount ratio in identifying different choice of cut-off and there are also some different observations to be noted and there are also several aspects in solving and this will play important role in the future work, this will be very helpful in bringing significant changes and development [1].

### *2.2 Measuring Consumption and Wealth Remotely Using Satellite Technology*

Night time light is hazardous sometimes makes it difficult to measure detection of amount of poverty, combining day time and night time images will make more easier way for conclusion of the detection. In this model some of the effective approaches such as the accuracy, scalability, inexpensiveness, all these factors will be used to determine the level of economic rate in a particular region. Based on the images the machine can predict the wealth and consumption of that region this makes it economical in determining the growth rate, and shows a statistical growth rate developed in any region that will be provided as input.

According to a research ,a survey was conducted based on the survey that was conducted the outcome derived from the survey and satellite data from an African countries such as Tanzania, Uganda, Malawi, Rwanda , it shows how a convolutional neural network was used to train the model , and how well was the machine able to get trained from the satellite data or the images data that is being passed , by the help of this type of data the machine can follow an ordered way of learning .

This then demonstrates us how well the machine learning techniques can be used to train the model in obtaining original implementation and its accuracy of providing output, this becomes an application for research for various kinds of domains such as scientific domains, public domains etc. Thus, accuracy of the output depends on the number of images provided as the output and these number of images can be any data, that is it can be the Night time data or Sun light data [2].

### *2.3 Poverty Mapping in Belize using open satellite and machine learning*

There are several methods proposed till now that have been proposed to determine the model poverty using satellite imagery and these methods outcome derived to be true and sometimes false, also every method generated outcome does not come to be true in which some of them also evaluated to false , this may be because of the convolutional network learning algorithm applied to the method applied directly or indirectly **,**But however these methods require good approaches and are found to be expensive and also cause infrequent acquisition.

Many scientific researchers have found different methods to solve these kinds of problems and hence there is no one specific approach to solve this problems ,many different algorithms have been designed to overcome such issues , it completely depends upon the manual needs and there measurement of accuracy they need , the type of accurate solution the model provides , this is necessary because each region might have different complexities and different behaviours of the way it can be used to predict the output, this algorithm in Belize used by including these satellite features and hence they came to a conclude that the open data is necessary for development and they must also have the pipelines of data open where possible[3].

### *2.4 Intra-Urban Poverty Measuring Using Data from Remote Centres*

This section improvises the importance of the remote sensing data and it shows the evaluability of the remote data centres and this methods are used to quantify the degree of the poverty and it helps to come to an conclusion to improve the level of understanding the degree of poverty this concept will be helpful for understanding the poverty level and improving the easiness of the level of complexity in finding the data of poverty from manually reaching out to every region, place, state, or country to find the number of people living there, and hence remote sensing of imagery will be helpful for a higher range when it comes to intra urban measurement scale , by this method the researchers can come to a conclusion that the physical condition or the appearance or physical state of the areas of the urban will the reflection of that society residing in any region, the state of any region its physical state define the whole society , based on its certain conditions such as the amount of agricultural activities involved in the region , the development of the road phase in it , industrial activities taking place in them and also the economical state of that particular region .

The second methods is the physical conditions of their house ,the house roofs the entrance of any house , house doors and many more such appearances will define the poverty level of that region ,hence its necessary to consider the housing conditions as one of the main factors of poverty predictions and hence this also makes it important for researchers to improvise and consider this as an important factor of considering when understanding the level of poverty and its ratio of rating for economical factor also , in this method one of the most high clarity pixel of images are used to determine the socially economically improved version of classification .

One of the another aspect to be considered is the that the world is filled with urban more than the rural areas in the whole world the amount of population is higher in the urban areas compared with the rural areas ,in every year there is a rapid increase in the population development in the urban areas and decrease in the rural areas ,so there is very high growth in the cities in every regions ,generally in many areas the growth of cities increases and they exceed the capacity of handling the government to provide any other services and any other kinds of infrastructures

The census data shows that the complexity in predicting the poverty areas becomes difficult in urban areas [4].

### *2.5 E-commerce Data Based Poverty Level Prediction*

There are many agencies that are concerned with finding out the level of poverty in each region and also improvise the focuses on concerning the economic and other household activities and many other social needs that are necessary and it makes an impact on the other kinds of activities that are involved and .For every two years the government conducts NSES which abbreviates as National Socio-Economic Survey and by this every year the government conducts different activities and sessions and based on the outputs or the decisions derived from it are to measure the poverty levels in the Many countries and also handle them , and by this the government can come to conclusion based on many other features they find out from that activities and also help for many other development methods and hence improvise them for handling the certain factors that cause the issue.

In this method the K-nearest algorithm is the mostly used algorithm and it has good factor of finding out or figuring out the necessary details of improvement and also help in the complement the census of BPS, this method has good accurate results of predicting the poverty levels and it can also provide the necessary information about handling or taking precautions when measuring the poverty levels at each region [5].

### *2.6 Costa Rican Household Poverty Level Prediction In R*

According to the report details of the current process of answering many of the problems related to the poverty levels and also the Costa and other Rican problems of the household related issue , it has come to an compromise that the data sources, the data analysis for exploratory through many of the factors

of visualization ,development of many of the models and also the different approaches that are involved in the disproportion glitches, and all other metrics of performances , and at the end gives the visualization results .And in a society where there is scarcity of many problems it is the importance for government to consider those regions and provide them with the necessary aid accordingly and also make them with the welfare social programs. And it was also observed or proved that in the backward areas the people are not able to provide with the necessary documents that are necessary also make them with the available sources of expenses or income records to prove the qualify for the aid.

In many other countries the popular method called as the Proxy Means Test, is used to arrive at a certain conclusion, there are many other factors based on which the company can decide the eligibility for poverty level predictions those are the number of rooms, the number of people in the household to conclude about a family qualification and all other details for aid. But though having many kinds of algorithms one of the major issues to be considered will be the accuracy and it can remain as a problem when predicting the poverty at kinds of level though using complex algorithms and these methods are being used to be the most used such as the convolutional neural network, k nearest algorithms, recurrent neural networks etc. The data source is extracted from Kaggle [6].

### *2.7 Random Forest Multivariate and Malnutrition Prevalence*

The recent advancements in the areas of the remote sensing have been increasing rapidly and also the area of remote sensing has good factors of providing the methods of generating the accuracy and also it has been considered to be the effective approach out of all the other approaches and this is also considered to be the accurate and inexpensive way of solving the problems and also make it effective for solving problems of any other regions also and the estimation that takes place timely and also told to be as the indicators for malnutrition and for the growth of development ,and however the state of the art models that often get relied upon are also considered to be as the important factor that represents the challenging factor to interpret .the malnutrition and the prevalence of the poverty are used to open access and also get geo referencing data. this makes it important to consider these kinds of methods.

The government and all other agencies are working very hard to figure out all the problems and in order to bring accuracy in such kinds of problems to find and resolve them ,they have devoted many of the considerable factors to improvise the poverty towards the reduction efforts of the malnutrition , and also using the current observations of the baselines features of the company have been found out to be geographic needs and also implement the past observations of the past records that record them in the baseline measures and for the evaluation of impact interventions . also, the contemporaneous predictions are made to inform all the incoming and outgoing activities evaluated. The more the precise the applications are the more inexpensive the data demands for the model, and also the greater the likelihood and also the agencies are made to be the accurately and evaluate interventions [7].

### *2.8 Poverty Level Characterization Using Feature Selection and Machine Learning*

Poverty is considered to be one of the most persistent and also the social and cultural problems that define the orders to develop all the necessary accessories and interventions problems are being found out policies. But the poverty wealthiest scale is not at all used as the way to categorize people and many other surveys that are conducted and also it helps to establish these kinds of this, and subjective opinions are frequently skewed out be as one of the options for data availability.

Classifying or predicting the poverty is one of the difficult approaches methods to solve and time consuming, this might be because of the lack of resources such as the security and data scarcity and also achieving these kinds of accuracy is also a major issue and some of the categories in these are the ways in which all the methods and creating an emergency fund. The first categorical problems are effective [8].

### 2.*9 Poverty Predictions And Mobile phones Meta Data*

In some of the developing countries, the wealth and the income expenditure calculation are a time-consuming process, where in it becomes difficult for the authority to come to a decision of calculating the amount of accuracy involved in it. And taking advantage of the omnipresence of the mobile phones in some of the cities such as Rwanda, alpenstocks etc.

They have applied many algorithms and many models in the development phase of this kinds of tasks and also approach towards all problems that they faced and it also has an impact and show that the matching of all the predictions happened well and are matched well with all the other data to be noticed. Since the problems arise with the accuracy different algorithms must be used in order to solve them [9].

### 2.10 *Three Level Multidimensional Poverty Prediction Using Machine Learning*

Poverty is considered to be ass one of the multidimensional notions which is not only related to one particular aspect but also to many aspects such's as considering some of the components such as the health status, consumption and social and context deprivation.in these types of works a good dataset is always necessary to find out the number of tasks needed to be accomplished.

This is one of the most significant problems to be found where in the agenda of 2023 defines the sustainable and development approved by the united nations and it also makes it inevitable process to end the level of poverty in all the levels and it also has been the major significant problem, in order to correctly identify the poverty risks involved in poverty and also makes it difficult approaches to solve them, some of the major difficulties are finding the labelled data is found to be as the complex problems and the finding the accurate datasets has also been the way of finding the effective way of selecting all the robust features to feed all the machine learning and deep learning algorithms ,to figure out the accurate result ,the right type of datasets is needed, and how different features will try to contribute to the project also puts impact ,and how it helps when using heterogenous data[10]

## 3. Conclusion

This work progress the type of work it handle even on heavy workload different type of datasets, it makes and informates that there are different ways that are been found out to solve the problem and the above one is a type of way for the problem to be solved , Thus the machine learning projects can be used to solve the problems in it and also make use of the remote sensing data and many other kinds of household survey data, thus this project conveys

datasets are mandatory for the classifications and accuracy to happen correctly and effectively. based on many factors this model predicts region to region poverty level by the help of complex algorithms.

**REFERENCES**


1. Gary R.Watmough, Peter M.Atkinson,Arupjyoti Saikia,Craig W.Hutton , "Multidimensional poverty and counting poverty measurement " , //doi.org/10.1016/j.worlddev.2015.10.031,volume 78, 2016, Science Direct.
2. Neal Jean, Marshall Burke, Michael XIE, W. Matthew Davis, David B.Lobell, " Measuring Consumption and Wealth Remotely using satellite technology", DOI: 10.1126/science.aaf7894,volume 353,2016,Science.
3. Jonathan Hersh, Ryan Engstrom, Michael Mann Criminal, "Poverty Mapping in Belize Using Open Satellite and Machine Learning" , //doi.org/10.1080/02681102.2020.1811945, Volume 27,2021, Issue 2.
4. Juan C.Duque, Jorge E.Patino, Luis A.Ruiz, "Intra-Urban Poverty Measuring Using Data From Remote Centers", LandScape and Urban Planning,2015 , Science Direct.
5. Tiara Fatehana Aulia, Dedy Rahman Wijaya, Elis Hernawati, Wahyu Hidayat, "E-commerce Data Based Poverty Level Prediction", 2020, IEEE.
6. Chaitanya Pramodh Kasula, Aishwarya Varala, "Intra-Urban Poverty Measuring Using Data from Remote Centers", TowardsData Science.com,2019, Towards Data Science.
7. Chris Browne,David S.Matteson,Linden McBride,Leiqui Hu, "Random Forest Multivariate And Malnutrition Prevalance",2021,PLOS ONE.
8. Mahek Sabha, Soundarya R, Spoorty Balaji, Niriksha, Ankitha Shetty," Poverty Level Characterization Using Feature Selection and Machine Learning", volume 9, 2021, ISSN:2320-2822, IJCRT.
9. Joshua Blumenstock,Gabriel Cadamuro,Robert On, "Poverty Predictions and Mobile Phones Metadata",DOI: 10.1126/science.aac4420,volume 350, 2015,Science.
10. Fabio D'Adda, Marco Cremaschi, Enza Messina, Marco Terraneo, Stefania Bandini, "Three level Multidimensional Poverty Prediction Using Machine Learning",2023, ITAL.AI.