Chapter 4
Machine Learning and Artificial Intelligence: Rural Development Analysis Using Satellite Image Processing

Anupama Hoskoppa Sundaramurthy
BMS Institute of Technology and Management, India

Nitya Raviprakash
Rashtreeya Vidyalaya College of Engineering, India

Divija Devarla
Rashtreeya Vidyalaya College of Engineering, India

Asmitha Rathis
Rashtreeya Vidyalaya College of Engineering, India

ABSTRACT

This chapter proposes a cost-effective and scalable approach to obtain information on the current living standards and development in rural areas across India. The model utilizes a CNN to analyze satellite images of an area and predict its land type and level of development. A decision tree classifies a region as rural or urban based on the analysis. A summary describing the area is generated from inferences made on the recorded statistics. The CNN is able to predict the land and development distribution with an accuracy of 95.1%. The decision tree predicts rural areas with a precision of 99.6% and recall of 88.9%. The statistics obtained for a dataset of more than 1000 villages in India are cross-validated against the Census of India 2011 data. The proposed technique is in contrast to traditional door-to-door surveying methods as the information retrieved is relevant and obtained without human intervention. Hence, it can aid efforts in tracking poverty at a finer level and provide insight on improving the economic livelihood in rural areas.

DOI: 10.4018/978-1-5225-9687-5.ch004

INTRODUCTION

Around one-fifth of the world's population is afflicted by poverty. India has one of the fastest growing economies in the world but severe poverty is prevalent across many parts of the country. According to the World Bank, as of 2012, 20% of the Indian population is poor out of which 80% is found in rural areas (Zhou, Yang, & Yu, 2013). Data released by the Socio- Economic Caste Census Survey of 2013 reveals that 75% of the rural population or 133.5 million families earn less than Rs.5000 per month (Tewari, 2015). Asset ownership is also reflective of the poor standards of living in such areas. Many houses are kutcha (made from low quality materials) and households do not have refrigerators or motorcycles. In many areas, basic facilities for education, sanitation, health, infrastructure, and transportation are lacking. The Ministry of Statistics stated that in 2013, 22.3% of males and 47.5% of females were illiterate in rural India. (DataGovIn, 2014). Due to a lack of proper education, there is insufficient skilled labor resulting in unemployment in rural areas (Cohen & Medioni, 1999). Rural India relies largely on agriculture, a low-income economy, as a means of livelihood. In addition to the many geographical factors that erratically influence this industry, there is a lack of proper road and railway transport in many localities thus destabilizing agricultural marketing (Jadhav, 2014). There is an urgent need to address sanitation issues such as improper sewage systems and shortage of toilets as well as health deficiencies like malnutrition and stunted growth in rural areas. Most of the people must travel more than 100 km just to avail a medical facility. The Ministry of Health and Family Welfare has stated that although the number of Primary Healthcare Centres (PHCs) and Sub Centres (SCs) have increased rapidly from 2005 to 2015, it still falls short by 5.21% of the overall requirement. (MHFW, 2015). It does not meet the expected standards of the World Health Organization with respect to the population norm (Shafie, Hafiz, & Ali, 2009). Many schemes are in place to propel the development of rural India and this can be advanced further by improving the quality and quantity of information regarding the problems faced by the poverty-ridden population. Measures of economic and living standards of populations influence both research and policy in a critical manner. Understanding these measures leads to informed policy making by governments such as the provision of fundamental services like water and healthcare. Although the economic data available in developing countries has improved in recent years, reliable data collection for poverty analysis is problematic. (DFID, n. d.). Manual surveying techniques are time-consuming, involve a huge expenditure of resources, and are difficult to conduct on a regular basis. For example, Census of India is only carried out once in ten years (MHA, n. d.). The complete set of statistics for every locality in a district is unavailable. For rural areas, systematic approaches for data collection is an arduous task as many regions, being remotely located, are often neglected. Often, the results are summarized at the district level making it difficult to target specific areas under poverty.

In 2015, the United Nations declared that the first Sustainable Development Goal is to eliminate poverty by 2030.(UNDESA, 2018). The main challenge behind achieving this goal is identifying all the regions that fall below the poverty line. If the exact locations are determined, it will help the government and nongovernmental organizations (NGOs) allocate the necessary resources accordingly, to improve the economic well-being of these rural areas. Using machine learning on satellite imagery, this paper proposes a technique to identify rural areas and help track poverty in India. It additionally provides a technique that reduces the manual effort and resources required in obtaining development-based statistics for rural localities.