

# Predicting Poverty Index using Deep Learning on Remote Sensing and Household Data

Parth Agarwal, Nandishwar Garg, Pratibha Singh



**Abstract:** The main challenge for approving and implementing policies aiming at sustainable development of country is correct prediction of socioeconomic condition. Deep learning algorithms in recent researches have been identified as potential resource to be applied in this domain. Another challenge is availability of sufficient amount of data which is solved using transfer learning in Convolutional Neural Network (CNN). We used pre-trained Inception Net-v3 and Ridge regression model to estimate poverty level using publicly available dataset comprising of daylight images, nightlight images and survey data. Each cluster of samples contains households between 1- 28. Its mean is 21.09, median 21 and a standard deviation is 1.36. Proposed deep learning inspired model estimates wealth-score for 28393 clusters with an  $r$  value i.e. Pearson Correlation Coefficient of 0.73, signifying  $r^2$  value i.e. Coefficient of determination of 0.54. It shows that daytime satellite images, nightlight intensity and demographic data available can be utilized for precise evaluations about the spatial scattering of monetary prosperity crosswise over different nations.

**Index Terms:** Convolutional Neural Network, Daylight, Nightlight, Regression, Satellite Images, Survey Data

## I. INTRODUCTION

The key issue is to estimate regional poverty level to make strategies for eradicating poverty. Precise estimations of the economic representation of community impacts both strategy and research. Our approach requires just publicly accessible data and may be able to change efforts for tracking and targeting poverty in developing nations. Precise estimations of the economic representation of community impacts both strategy and research. Such estimations frame choices by governments about how to apportion limited assets and establish a support to worldwide attempts to comprehend and follow progress towards improving human livelihoods. Solid information on monetary jobs stay rare in creating scene, hampering endeavors to examine these results and to plan arrangements for any improvement. Here in this paper, we validate an exact, inexpensive, and scalable method for evaluating utilization of resources from high-resolution satellite images. We show how transfer

learning in a convolutional neural network can be instigated to recognize image features from relatively smaller dataset. Proposed model, which requires only publicly available data, could change endeavors to estimate poverty level in developing nations.

The issues that can be handled using our model are as:

- The predictions that our model provides can be used by policy makers for building policies to eradicate poverty.
- Expenditures and asset wealth can be estimated for those districts scaled up to countries where any survey data is not available.
- Our trained model can be used for countries for which it has not been trained on and sufficient data is not available for training.
- We can help to bridge the gap to recognize variation in data and find those areas effectively with utmost need.

## II. LITERATURE SURVEY

In developing world reliable information on economic livelihoods remain scarce that can help to predict socioeconomic parameters of a country and plan policies that improve them. It encourages researchers to incorporate remote sensing data and use machine learning techniques for predicting important influential parameters. The suggested approach shows that current daytime satellite images of high resolution can be utilized to make perfect estimation about the spatial distribution of fiscal prosperity across nations.

Researchers in this paper [1] use CNN model on high resolution satellite images of five African countries to predict poverty. They transfer learning of pre-trained model of ImageNet on labelled images from 1000 distinct classes [2]. Next the CNN uses information learned from image classification problem and is fine tuned to predict the nightlight intensities provided daytime satellite images. These daytime images along with demographic health survey data is used to train ridge regression model in CNN that can estimate cluster-level score of wealth. They used daylight Images retrieved from Google Static Maps API [3], nightlight Images from DMSP-OLS Dataset [4] and survey Data from DHS Program [5]. Despite the fact that their model do better than models developed using other sources of passively collected data in predicting economic growth at the cluster level, they are as of now unable to measure variations within clusters as publicly available survey data allocate indistinguishable locations to all households in a given cluster to maintain respondent confidentiality. Proposed model can make predictions for daytime satellite images of any resolution which can be easily obtained, however predictions at finer level may probable be noisier [1].

Manuscript published on 30 September 2019

\* Correspondence Author

**Parth Agarwal\***, Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, U.P., India

**Nandishwar Garg**, Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, U.P., India

**Pratibha Singh**, Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, U.P., India

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

## Predicting poverty index using Deep Learning on remote sensing and household data

Another work has shown the prospect of advancing low resolution publicly available satellite images. Temporal model of CNN is used to predict the change in nightlights of four African countries across multiple years along with the prediction of wealth and consumption levels in a single year. Transfer learning approach is used to overcome the deficiency of data points at ground truth level. They used residual network model pretrained on the ImageNet dataset to estimate changes in poverty level across years using satellite images available through US Government's Landsat 7 satellite imagery program, nightlight intensities and publicly available LSMS poverty data. They used a ridge regression model with L2 regularization to avoid overfitting due to smaller dataset and calculated coefficient of variance  $r^2$  value [10].

Multilevel approach suggested to consider several other predictor variables covering more details of spatial landscape data. It has improved the estimation of household of Kenya by 10% interestingly by examining relationship between household and remote sensing features as homestead, agricultural land, village clusters and periphery [12].

In sequence of a series of such contributions, one is to identify the location for the informal settlement of vulnerable people using even low resolution images. They used back propagation algorithm with weights trained on PASCAL dataset and demonstrated classification schemes on their curated datasets [13].

The nightlight images with transfer learning on ImageNet

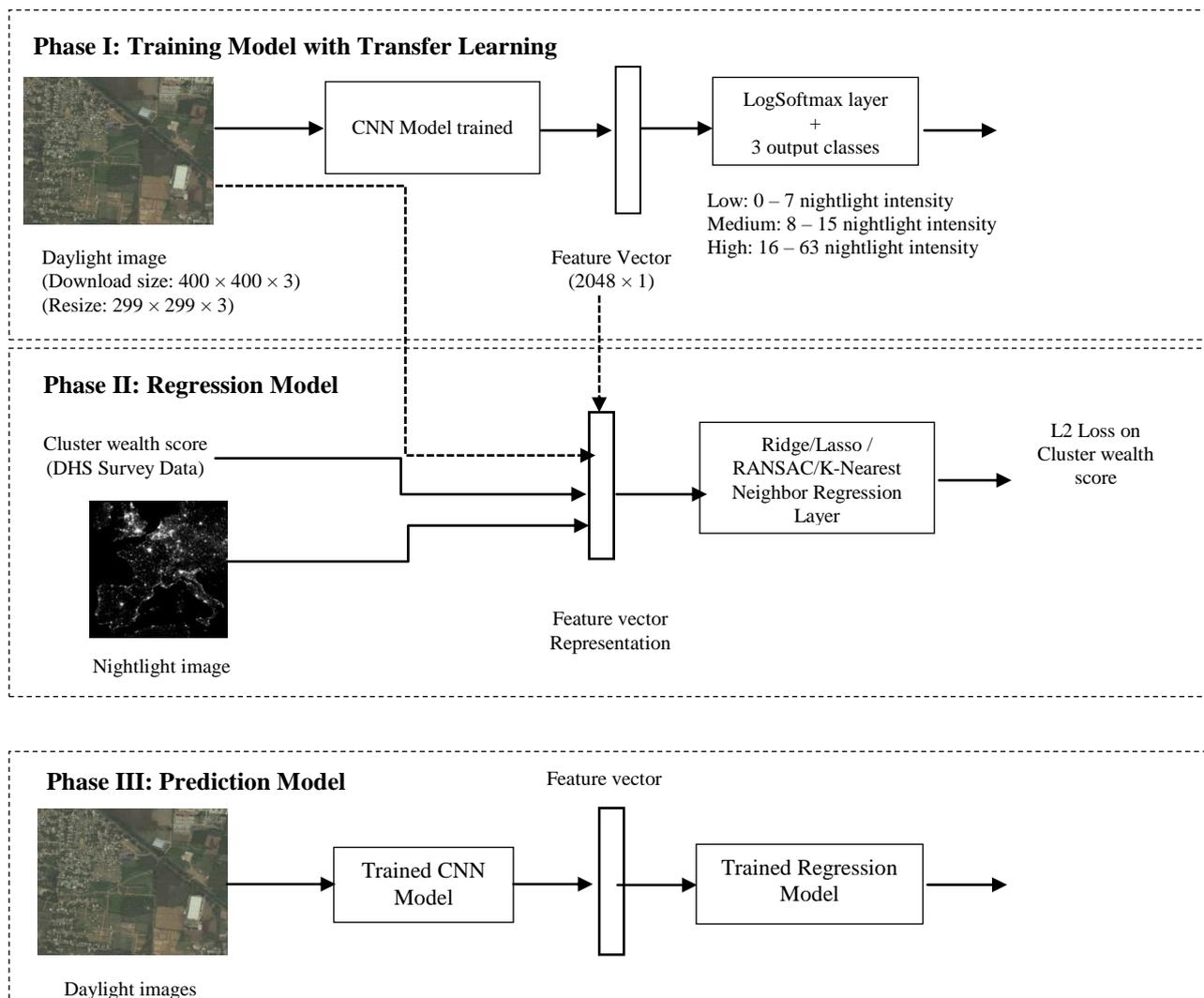


Fig. 1: Poverty prediction using Inception v3 and ridge regression

A deep CNN model was trained to simultaneously predict source of light, water and material of roof from satellite images to further estimate the poverty levels. It has used 2011 census data for statistical data to identify ground truth and Google static Maps API for extracting images for estimation of income level [11].

dataset is demonstrated for poverty mapping through high resolution satellite images of Uganda. The fully convolutional CNN model converts fully connected layers in VGG network to convolutional layer and is trained to identify landscapes and human constructed structures [14].

Broadly, only two type of sensors DMSP-OLS and SNPP-VIIRS are the sources for the night time light datasets. DMSP-OLS was firstly designed for spotting clouds at night but later on identified to be expedient for capturing radiance from city nightlights. We used DMSP-OLS dataset available in Geo TIFF format [15].

### III. PROPOSED MODEL

The three main phases of our proposed poverty prediction model are training of CNN model with transfer learning, training of regression model and poverty prediction model. These phases are clearly depicted in Fig 1 along with its main contributing components. These components are briefly described as below:

#### A. Daylight Images

Satellite Daylight images are downloaded from HERE Maps' Map image API. Each image is of dimension  $400 \times 400$  pixels [6]. This resource is used to download images corresponding to obtained geocoordinates (latitude and longitude).

#### B. Inception Net-v3

It has outperformed over Google Net, VGGNet, PReLU etc. architectures on the ILSVRC 2012 benchmark classification problem with much lesser computation cost. Inception-v3 Net is trained on ImageNet dataset which consists of 1000 classes [7] [8].

#### C. Nightlight intensities

Nighttime light intensities are extracted from nightlight satellite images. These nightlight images are extracted from time series data available for year from 1992 to 2014 on DMSP OLS data source. Dimension of downloaded image is  $43201 \times 16801$  pixels [4].

#### D. Trained CNN

We used CNN architecture of Inception Net-v3 trained on ImageNet dataset for predicting nightlight intensity from daylight images. The number of neurons in last layer has three output neurons corresponding to 3 output classes (Low, Medium, and High). Nightlight intensity is a proximate measure for economy of an area to a certain extent. Features extracted from the last convolutional layer (Mixed\_7c) of an Inception module which provides feature tensor of  $2048 \times 8 \times 8$  dimension. Each image after average pooling produces 2048-dimensional feature vector.

#### E. DHS Survey data

Demographic Health Survey data of India is used to extract cluster wealth score. The survey locations are used to extract samples for training for further prediction of poverty level [5].

#### F. Cluster wealth score

Wealth index is computed as the main component of responses of inspection of commodities. A set of questions about ownership of common assets for example bicycles, televisions, materials used for house construction are investigated. These parameters are normalized for each country and therefore we do not require further normalization.

#### G. Regression Model

Ridge Regression model also known as Tikhonov regularization is trained to predict cluster wealth score for the features extracted from CNN for corresponding daylight images of the cluster. This model solves a regression model where loss is calculated using linear least squares function

and L2 norm regularization is used to avoid overfitting.

#### H. Trained Regression Model

This regression model is further trained to predict a cluster wealth score for a feature vector generated by CNN model corresponding to a cluster's daylight image.

#### I. Trained CNN + Regression Model

Trained CNN and regression models are used further to predict cluster wealth score for a given daylight image.

### IV. IMPLEMENTATION AND RESULTS

Steps involved in implementation of proposed model are described below:

- a) Download Demographic Health Survey (DHS) from its program website [5].
- b) Download Nightlight Satellite images from DMSP-OLS website [4].
- c) Process DHS survey data to find parameters as latitude, longitude and wealthscore for households and group these household surveys into clusters.
- d) Perform geographical modifications on extracted locations from the survey data, to find values of geographical location (latitude, longitude) to download daylight images along with nightlight intensity corresponding to those coordinates.
- e) Download daylight images consistent to the locations obtained from previous step to generate a dataset comprising of 60,000 images classified into 64 categories based on the nightlight intensity.
- f) Initialize CNN model with the derived architecture of Inception-v3 and also weights learned through training of this Net on ImageNet dataset [8].
- g) Further training of CNN is executed on 60,000 daylight images. These images are classified into 3 classes with balanced data. The 0-7, 8-15 and 16-63 nightlight intensities are assigned to Low, Medium and High classes respectively.
- h) Parameters:
  - Architecture used is Inception-v3 with last layer's dimensions changed to (768, 3) with 3 outputs consistent Low, Medium and High classes
  - Image Size:  $299 * 299$
  - Hyperparameters:
    - Batch size = 96
    - Learning Rate (LR) = 0.001
    - Epochs = 15
    - Loss = Cross Entropy Loss
    - Optimizer = Stochastic Gradient Descent
    - Momentum = 0.9
    - LR Scheduler is StepLR with step size 8, gamma value is 0.1
  - Loss – 0.9614
  - Accuracy – 74.65%
- i) Extract features (Layer: 'Mixed\_7c') for all images in the dataset using the trained CNN.
- j) Use extracted features and cluster wealth score to train Ridge regression model after reducing dimensions to 1000 by PCA. We used 10-fold outer Cross Validation.
- k) The results obtained in terms of cross entropy loss and accuracy are shown in Fig 2 and Fig 3 respectively

The relation between our ridge regression model's predictions and actual cluster wealthscore values are shown in Fig. 3.

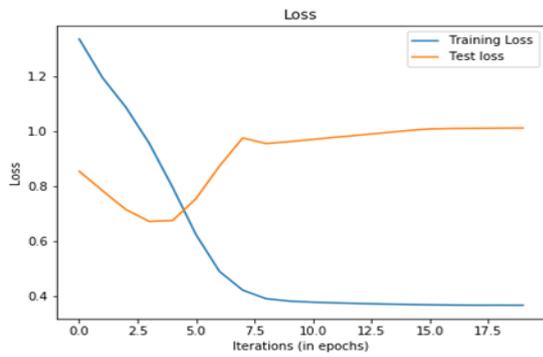


Fig. 2: Cross entropy loss vs iterations on nightlight intensities

As shown in Fig. 2, the training loss decreases swiftly from 1<sup>st</sup> to 7<sup>th</sup> iteration and drops slowly after 7<sup>th</sup> iteration because learning rate is decreasing gradually by a factor of 10. The test loss decreases initially but starts increasing from epoch 44<sup>th</sup> iteration which shows that the training has started to diverge at this point.

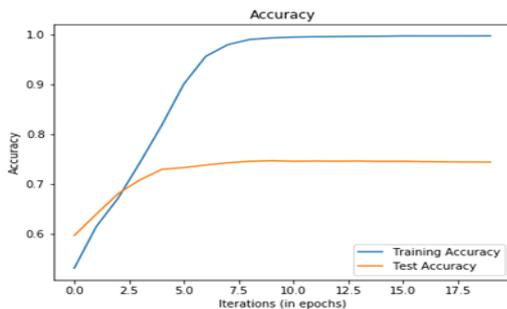


Fig. 3: Accuracy vs iterations on nightlight intensities

As shown in Fig. 3, the training accuracy increases rapidly from 1<sup>st</sup> to 7<sup>th</sup> iteration and after that it converges slowly to a saturation state because of gradually decreasing learning rate. The test accuracy increases rapidly till 8<sup>th</sup> iteration and gets saturated after this stage.

The results as shown in Fig. 4 depicts that our model's predictions are fairly accurate with a  $r$  value of 0.73 considering the issues that DHS cluster locations have noise mixed in them to preserve the anonymity of the survey and also the CNN has not been trained up to the optimum.

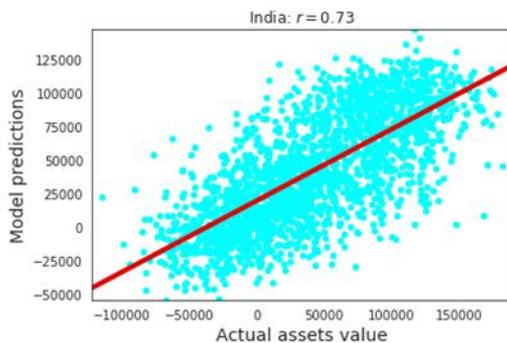


Fig. 4: Comparison between wealthscore of actual cluster (X-axis) and of predicted (Y-axis) cluster in terms of wealthscore ( $r$ )

Confusion matrices for three class labels of High, Medium and Low with and without normalization are shown in Fig. 5 (a) and 5 (b).

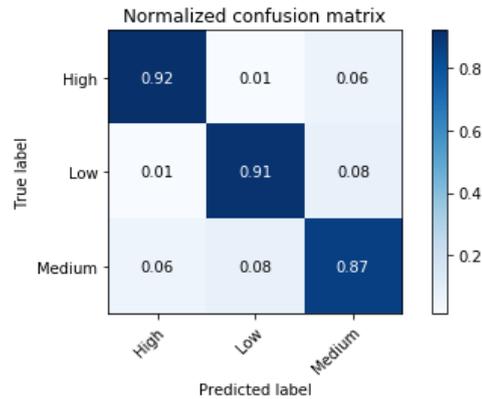


Fig. 5 (a): Confusion Matrices with normalization

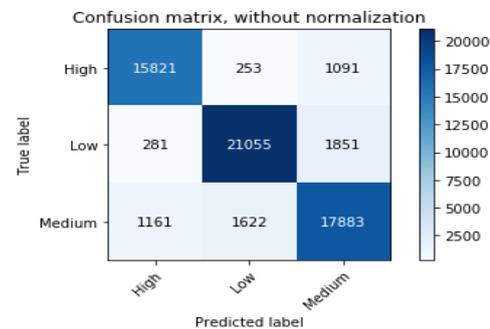


Fig. 5 (b): Confusion Matrices without normalization

Table I shows comparisons on correlation values of  $r^2$  on different regression models as Ridge, Lasso, RANSAC and K-nearest neighbors.

Regression models produce  $r^2$  value in the range of 0.5 to 0.54. Ridge regression with an alpha value of 10 achieves best  $r^2$  value of 0.533. Our model predicts wealth score for 28393 clusters having average households prediction with  $r$  value i.e. Pearson Correlation coefficient of 0.73 and  $r^2$  value i.e. coefficient of determination of 0.54.

Table I: Comparison of  $r^2$  value of different regression techniques with varying parameter values

Algorithms	Parameter value	$r^2$
Ridge Regression	Alpha = $10^{0.66}$	0.511
	$10^1$	0.533
	$10^{1.33}$	0.529
	$10^{1.66}$	0.524
Lasso Regression	Alpha = $10^{1.22}$	0.494
	$10^{1.66}$	0.523
	$10^{2.1}$	0.507
	$10^{2.54}$	0.472
RANSAC Regression	Base estimator = Ridge with alpha of 10	0.509
	Min Samples = 90%	

<b>K- Nearest Neighbors Regression</b>	Number of neighbors (K)	$r^2$
	2	0.401
	3	0.459
	4	0.484
	5	0.498
	6	0.507
	7	0.505

**Table II: Comparison of results**

Parameter	Jean et al. results [1]	Our results
$r^2$ (Coefficient of Determination)	0.68 (Nigeria), 0.57 (Tanzania), 0.69 (Uganda), 0.55 (Malawi), 0.75 (Rwanda)	0.54 (India)
Cluster size	3034 clusters	28393 clusters

As shown in Table II, our approach produces an  $r^2$  value of 0.54 which is comparable to that of Jean et al.'s results for Tanzania and Malawi. The relatively huge cluster size of our dataset for India is the main reason for us achieving this value of  $r^2$ .

### V. CONCLUSION

Predicting socioeconomic parameters (poverty index) using Machine Learning on remote sensing and household survey data in India will surely help in tracking poverty level, planning and implementing policies for further development. Primarily the proposed idea is inspired by the vision of Digital India, where every Indian is digitally endowed and every information is digitally available. Next and most importantly this idea is determined to play a key role for sustainable development as eliminating poverty has first priority in United Nations Sustainable Development Goals [9]. Our method, which requires only publicly available data (Satellite Daylight Images, nightlight satellite images produced by the DMSP OLS data source and DHS survey data) could change endeavors to track and target poverty in developing countries. Proposed work will estimate wealth score of any area using satellite imagery and survey data by applying methods of Artificial Intelligence. It will reduce the data gap and assist to understand poverty distribution throughout regions which is preliminary to build policy and allocate resources to reduce poverty levels. Possible extensions that can be made at the next stage of its implementation and uses are as following:

- Government can build policies for eradicating poverty
- Use different architectures for the CNN model or use Recurrent Neural networks (RNN) for better results
- Train model to the optimum with different architecture and hyperparameters and better hardware
- Use latest DHS datasets
- Use Google Static Maps API instead of HERE Maps
- Develop an android app/ website accessible for all naive users

### REFERENCES

1. Jean, Neal, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell, and Stefano Ermon. "Combining satellite imagery and machine learning to predict poverty." *Science* 353, no. 6301 (2016): 790-794.
2. Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al. "Imagenet large scale visual recognition challenge." *International journal of computer vision* 115, no. 3 (2015): 211-252.
3. Google Static Maps API, <https://cloud.google.com/maps-platform>
4. DMSP-OLS Dataset, <https://ngdc.noaa.gov/eog/dmsp/>
5. DHS program, <https://dhsprogram.com/data/available-datasets.cfm>
6. HERE Maps' Map Image API, <https://developer.here.com/documentation/map-image/topics/>
7. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2016) 2818–2826
8. InceptionNet V3 trained on ImageNet dataset, [https://download.pytorch.org/models/inception\\_v3\\_google-1a9a5a14.pth](https://download.pytorch.org/models/inception_v3_google-1a9a5a14.pth)
9. United Nations. Sustainable Development Goals: 17 Goals to Transform Our World. Available at: <http://www.un.org/sustainabledevelopment/sustainable-development-goals/>
10. Rastogi, Ruchir. "Temporal Poverty Prediction in Developing Countries", <http://cs229.stanford.edu/proj2017/final-reports/5244050.pdf>
11. Pandey, Shailesh M., Tushar Agarwal, and Narayanan C. Krishnan. "Multi-task deep learning for predicting poverty from satellite images." In Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
12. Watmough, Gary R., Charlotte LJ Marcinko, Clare Sullivan, Kevin Tschirhart, Patrick K. Mutuo, Cheryl A. Palm, and Jens-Christian Svenning. "Socioecologically informed use of remote sensing data to predict rural household poverty." Proceedings of the National Academy of Sciences 116, no. 4 (2019): 1213-1218.
13. Gram-Hansen, Bradley, Patrick Helber, Indhu Varatharajan, Faiza Azam, Alejandro Coca-Castro, Veronika Kopackova, and Piotr Bilinski. "Mapping Informal Settlements in Developing Countries using Machine Learning and Low Resolution Multi-spectral Data." arXiv preprint arXiv:1901.00861 (2019).
14. Xie, Michael, Neal Jean, Marshall Burke, David Lobell, and Stefano Ermon. "Transfer learning from deep features for remote sensing and poverty mapping." In Thirtieth AAAI Conference on Artificial Intelligence. 2016.
15. Jeswani, Reshma, "Evaluation of the consistency of DMSP-OLS and SNPP-VIIRS Night-time Light Datasets." (2017)

### AUTHORS PROFILE



**Parth Agarwal** has done B.Tech in Computer Science & Engineering from ABESEC Ghaziabad in 2019. He has secured good positions in e-Yantra, Google Developer Group's JIIT Dev Fest, Technotrix Coding Event, Quizzotica, Cognizance and is also Ex-Google Play Store Developer with 36,000+ installs.



**Nandishwar Garg** has done B.Tech in Computer Science & Engineering from ABESEC Ghaziabad in 2019. He has secured good positions in ICPC, TCS Codevita, Facebook HackerCup, E-Yantra, KPIT Sparkle, Vertigo Code Hunt and also received Silver Medal in B.Tech for second highest score.



**Pratibha Singh** is an Associate Professor in Department of Computer Science & Engineering in ABESEC Ghaziabad. She is B.Tech (HBTI K) and M.Tech (IITD). She is co-author of books on Operating System, Distributed System and Pattern Recognition and has published papers in International conferences

and journals.