IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

An Innovative Method to Calculate the Economic Development Index of Important Cities in West Bengal from Satellite Imagery

Tanuj Sur[1*], Asoke Nath[2]

[1]Dept. of Statistics, St. Xavier's College (Autonomous), Kolkata, India
[2]Dept. of Computer Science, St. Xavier's College (Autonomous), Kolkata, India

Corresponding Author: surtantheta@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i4.411417 | Available online at: www.ijcseonline.org

Accepted: 13/Apr/2019, Published: 30/Apr/2019

Abstract — Identifying the economic situation at sub-regional level has always been a challenge due to data constraints. In this paper we show the potential of remote sensing technologies to provide a possible solution to this problem. In this paper, we use the approach of using convolutional neural networks and transfer learning to process satellite imagery of the towns of West Bengal which can be used to predict its economic development indicator values.

Keywords — Deep learning, Convolutional Neural Networks, Transfer Learning, Asset Index, Satellite Imagery, Night Light Intensity, District Census Handbook.

I. INTRODUCTION

In this era of digitization and remote sensing technology, dependency on Census data is gradually decreasing. Platforms like Google Earth provide resources for geo-spatial analysis. One can write geo-scripts on Google Earth to extract data like Night Light images, daytime satellite images from various satellites. As a result of these technological advances we may no longer have to conduct surveys in the recent future to form an idea about the average financial situation of a town. Or at least we can use them to cross-validate the surveys conducted by Government or other organizations.

This paper aims to achieve this in 3 steps. *Firstly*, we use night light intensity data to extract informative features from the corresponding daytime satellite images using a CNN trained on ImageNet. In this case we use the night light intensities as predictors for wealth distribution in a town. We have observed an accuracy of 97% in classifying the daytime satellite images into low, medium, high nightlights. Our model trained on Nightlight intensity data in step 1 can clearly classify high level features like roads, agricultural lands, houses and water bodies. *Secondly*, we create an economic indicator from the District Census Handbook Data of West Bengal, 2011 and Google Nearby Search Api. The economic indicator is first PCA (Principal Component Analysis) of certain fields from these sources of data mentioned above. *Thirdly*, we train a ridge regression on the image feature vectors extracted from the CNN (being the predictor) and the economic indicators(being the predicted variable). This ridge regression predicts the economic indicator given any image feature vector.

The index predicted is used for comparative purpose to group towns in any state of India on the basis of certain threshold. This will help in identifying social structure based on wealth and resources. Also, the model calculates two indices, one based on 2011 Census data, and the other based on recent data available through Google Nearby Search api. The two indices can be used to compare the development of a town and nation as a whole in these two time periods.

II. RELATED WORK

CNN models trained on ImageNet are recognized as good generic feature extractors. Donahue et al. [1], show that features learned from training on the ImageNet dataset achieve world class results in object detection, classification. Satellite image classification depended mainly on human annotated data. This was a major disadvantage as labeled data was not readily available. However works by Jean et al. [2], have used night lights as proxies to solve this problem. Early works showed that places emitting high levels of artificial lights tended to have higher economic output. But nightlights have difficulty in differentiating between poor, densely populated areas. Also, nightlights can be misleading due to reflection of light from water bodies and shiny surfaces like desserts. Jean et al. used a two-step transfer learning model to estimate sub-regional levels of asset and poverty in five countries of sub-Saharan Africa. Andrew et al. [3], has measured human development from Satellite

Imagery. They have shown performance on wealth predictions for Nepal with R^2 =0.64 and for sub-Saharan African nations with R^2 =0.51. They have used Demographic Health Surveys (DHS) as a source of ground truth for development outcomes.

Noor et al. [4] showed that nighttime luminosity correlates with asset based measures of wealth in 37 African countries. Much of the work on wealth and human development prediction has been done in the African countries due to the lack of structured and reliable data available in that continent. However in countries like India where structured data on household and infrastructure is available (in the form of Census data and others) , major challenge is the duration of the cycle after which the surveys are conducted. In India, Census data is published once in every 10 years. In 10 years, the financial market can experience lots of variations and there can be notable changes in the shift of economic strata in the society.

III. METHODOLOGY

We elaborate the methodology sequentially as mentioned in the steps of Introduction. *First* we trained our CNN on night light intensity data to predict categories of high, medium, low from daytime satellite images of various places of India. This CNN is pre-trained on ImageNet. This model learnt all the high level features and it can identify houses, greenery, fields. In the *second* step we define an economic development index calculated from the fields of District Census Handbook for Towns of West Bengal, 2011 and Google Nearby Search api. *Thirdly*, we train a ridge regression on the image features and economic development index. This project is implemented for towns of West Bengal. However it can be implemented for all States of India.

A. COLLECTION AND PREPARATION OF DATA FROM IMAGENET AND VIIRS DATASET

The training of CNN on nighttime light intensity data involves collection of data from VIIRS and using a model pre-trained on ImageNet.

We started with a CNN pre-trained on ImageNet which can identify objects from the ImageNet dataset. This is done so that the CNN has prior knowledge about low level features beforehand like edges, corners, curves. The ImageNet dataset is a large image classification data set that consists of labeled images from 1000 different categories. The network pre-trained on the ImageNet is fine-tuned using satellite images and nighttime light intensity data through transfer learning. Transfer learning refers to the passing or transfer of knowledge learnt from some field into another field. In this case, the weights of the ImageNet dataset are used for further task of classification.

For twenty years, till 2011, DMSP satellite had collected images of the nighttime throughout the world for every night. However since we are training on images of 2018, it was necessary to use the nighttime intensity data collected during this period. We obtained the data of 2018 from VIIRS satellite. This was done by geo-scripting on Google Earth, corresponding to the coordinates of places in the images we trained our model on. So that the pre-trained model learns all kinds of high level visual features like roads, vegetation, water bodies, houses, fields etc., we train the model on images collected from all over India. The CNN is trained on 375,000 daytime satellite images. These satellite images are collected through Map Box API [6]. The dimension of each image is 512x512 pixels.

After collecting the VIIRS data on night light intensities, the average value of night light intensity data is then considered i.e. average of all monthly values. Following Jean et al. we then convert each value into one of the three classes: low, medium, high. We assigned the low intensity category to pixel values from 0 to 2, the medium intensity category to pixel values 3 to 34 and high intensity category pixel values from 35 to 70. The image id refers to the name of the images, stored in the database, used for training our model. For example, 4 such images/locations are mentioned in the Table 1.

The nighttime light intensities are used as training labels after they have been categorized into high, medium or low intensities. It is made sure that we have equal number of images for each category before training the CNN on it. This is done by sampling randomly equal number of images from each class. In this case each class contains 100,000 images. Examples of image from each class are given below.

TABLE 1: MONTHLY VIIRS DATA OF 2018 FOR 4 LOCATIONS

IMG_ID	January	February	March	April	May	August	September	October	November	AVG
97701_57110	58.22	54.64	41.1	51.26	51.39	52.8	45.81	86.44	65.34	56.33
97701_57112	49.03	49.95	39.8	46.51	44.09	52.08	37.97	97.61	66.2	53.69333
97701_57114	49.17	49.66	39.17	44.82	41.05	47.54	38.58	104.2	65	53.24333
97701_57116	53.65	53.37	40.07	48.22	42.62	48.51	39.54	112.43	67.55	56.21778

Figure 1. High Intensity

Figure 2. Medium Intensity

Figure 3. Low Intensity

B. TRAINING OF THE MODEL

Image augmentation is necessary to increase the training data size. Hence the images used for fine tuning the CNN are horizontally flipped and rotated by 0.2 degrees inside the Image Data generator. Here we have used Resnet50 architecture. Size of each image is reduced to 224x224 pixels and batch size is 64. Choice of batch size is important because a very large value reduces the networks ability to learn the features in depth resulting in under-fitting. Similarly very small value may cause over-fitting as pointed by Nitish Shirish Keska et al. [5]. Then, we removed the fully-connected top layers from the CNN. At the top we added dense convolutional layer with 1024 hidden units and activation relu followed by global average pooling layer, and dense convolutional layer with 3 hidden units and activation sigmoid. The training set contains 80% of the images while the validation set contains 20% of the data. Each town had around 140 to 156 training images covering an area of 10 square kilometers. The network architecture is as follows:

INPUT→[[CONV→RELU]→POOL?]→[FC→SIGMOID]

The last layer performs the classification task hence sigmoid activation function is used. The sigmoid output gives 3 values between 0 and 1. The highest value represents the category to which the image belongs. Optimizer used is rmsprop and loss function is categorical cross entropy. The learning rate is 10^{-2}. The model is trained for 10 rounds of 5 epochs. After first 5 rounds, first 140 layers of the network were frozen and the remaining network was trained. This ensured that the newer untrained layers are fine-tuned so that they can learn high level features. The lower layers contribute in identifying low-level features hence not much training is required with the satellite images. The final training summary is given below.

```
Epoch 1/5
51/51 [==============================] - 54s
1s/step - loss: 0.1287 - accuracy: 0.9567
Epoch 2/5
51/51 [==============================] - 53s
1s/step - loss: 0.0936 - accuracy: 0.9678
Epoch 3/5
51/51 [==============================] - 53s
1s/step - loss: 0.0870 - accuracy: 0.9679
Epoch 4/5
51/51 [==============================] - 52s
1s/step - loss: 0.0680 - accuracy: 0.9735
Epoch 5/5
51/51 [==============================] - 52s
1s/step - loss: 0.0679 - accuracy: 0.9743
CPU times: user 7min 39s, sys: 21.7 s, total: 8min
Wall time: 4min 24s
```

As we can see, the loss has gradually gone down and the accuracy has increased. The final accuracy is 97.43 %. This accuracy is based on k-fold cross validation. After tuning the Convolutional neural Network, it can be thought as a function whose input is images and output is 1024 dimensional vector of activations in the top layer.

C. CONSTRUCTING ECONOMIC DEVELOPMENT INDICATOR

The economic development indicators represent the financial situation of a town. We calculate the economic indicators for both 2011 and 2018. The indicator for 2011 is based on District census Handbook for Towns of West Bengal,2011(DCHB) and the same for 2018 is calculated from Goggle Nearby search api. The fields used to compute the indicator in both cases have causal relation with wealth distribution of a town.

D. ECONOMIC DEVELOPMENT INDICATOR FROM DCHB DATA, 2011

The DCHB data contains 429 columns. To measure the economic index at town level, we use the fields Nationalized Bank (Numbers), Private Commercial Bank (Numbers), Co-operative Bank (Numbers), Agricultural Credit Society (Numbers), Non-Agricultural Credit Society (Numbers). The fields represent the economic activity of a town. This rationale was given by Ghosh and De(1998) [7] who argued that number of banks in a region is positively correlated to the cash inflow and outflow from a town. Since, more the economic activity of a town, more is its asset valuation hence these fields were chosen as an indicator for economic development. Once the data is cleaned, the columns are standardized and Principal component analysis is applied which reduces the multi-dimensional data to a single a dimension i.e. from 5 dimensions to a single dimension. There is always some loss of data due to PCA. In this case we retain 56.4 % of the variance. Table 2 shows the standardized set of 5 towns and asset indices of 10 towns of West Bengal. The values in Table 3 are indeed representative of the economic development. Kolkata seems to have the highest value which shows it is economically most developed among rest of these places followed by Siliguri, Durgapur, Haldia, etc. Hence this asset index can be used for comparison of the economic activity of towns.

TABLE 2: VALUES OF THE FIELDS FOR 5 TOWNS IN WEST BENGAL

Town	Nationalized Bank	Private Commercial Bank	Co-operative Bank	Agricultural Credit Society	Non-Agricultural Credit Society
Asansol	-0.247590743	-0.269866001	0.298142397	-0.616669839	-0.606052032
Balurghat	-0.321566026	-0.283743408	0.298142397	-0.616669839	-0.606052032
Barasat	-0.325088659	-0.276804704	0.298142397	0.342594355	-0.606052032
Bardhaman	-0.208841785	-0.280274056	2.385139176	3.220386936	-0.595709006
Darjiling	-0.325088659	-0.280274056	-0.745355992	-0.616669839	-0.506069453

TABLE 3: TOP 10 ASSET INDEX (PCA VALUES) FROM TOWNS OF WEST BENGAL

Town Name	Asset Index
Kolkata	18.34619788
Siliguri	2.88052899
Durgapur	1.973444739
Haldia (M)	1.065703685
Tarakeswar (M)	1.006896123
Bardhaman	0.794525755
Tamluk (M)	0.397785135
Kanchrapara (M + OG)	0.227595959
Kalimpong	0.198805096
Baranagar (M)	0.151724389

E. ECONOMIC DEVELOPMENT INDICATOR FROM GOOGLE NEARBY SEARCH

Here, we call Google Nearby Search Api to extract the data. The advantage of this method is that, the data is recent and does not depend on Census or any other sources of data. Nearby Search lets you search for places within a specified area. You can refine your search request by supplying keywords or specifying the type of place you are searching for. In this case the fields we considered were number of atms, number of banks, number of supermarkets, number of car dealers. As mentioned above, banks and atms are indicators of cash flow, super markets and car dealers also represent the financial status of a town. More number of super markets shows ability of people to shop. Similarly more number of car dealers in a town indicates that town must be wealthy. These columns are standardized and Principal component analysis is applied which reduces the multi-dimensional data to a single a dimension i.e. from 4 dimensions to a single dimension. There is always some loss of data due to PCA. In this case we retain 73 % of the variance. This shows there is less loss in data compared to the previous method. The tables 4 and 5 show the standardized set of 5 towns and asset indices of 9 towns of West Bengal.

The asset indices calculated from the two methods should not be compared as these two indices need to be standardized so that they are on the same footing.

TABLE 4: VALUES OF THE FIELDS FROM NEARBY SEARCH FOR 5 TOWNS IN WEST BENGAL

Town Name	No. of atm	No. of bank	No. of supermarket	No. of car dealers
Amkula (CT)	-1.0555253	-1.2072296	-0.649219919	-0.164967474
Badamtam Tea Garden (CT)	-0.6365206	-0.1584258	-0.373244901	-0.164967474
Baranagar (M)	1.7378388	1.63952341	-0.097269883	2.460167106

Barrackpur Cantonment (CB)	1.1791660	1.6395234	-0.097269883	1.001759006
Barunda (CT)	-0.9158571	-0.7577422	0.178705135	-0.456649094

TABLE 5: TOP 9 ASSET INDEX (PCA VALUES) FROM TOWNS OF WEST BENGAL

Town Name	Asset Index
Kolkata	6.542982
Barasat	6.404887
Bardhaman	3.796285
Purulia	3.670349
Asansol	3.406318
Baranagar (M)	2.866098
Bally (CT)	2.853939
Chandannagar (M Corp)	2.715844
Barrackpore (M)	2.577749

F. FITTING OF RIDGE REGRESSION

Through the steps discussed above, we have a CNN that can produce 1024 dimensional feature vectors from each satellite image and the asset index calculated in either way. We combine these two by first creating a pool of all satellite images surrounding a town covering an area of 10 sq. km as done by Andrew et al. Then each image is converted into its corresponding feature vector by the method mentioned in 3.1.2. All the images in the pool for a town is averaged into a single feature vector for each town. This is our predictor variable. Finally we use mean town level image features extracted from daytime satellite imagery and the corresponding asset index to train ridge regression model that can estimate asset index given a daytime satellite image. The feature vector is a higher dimensional tensor, that's why we use multiple regression(here ridge regression) to train this model. We generated the daytime satellite images for 480 towns in West Bengal by creating a bounding box of area of 10 sq. km. with the centre being the coordinates of the town. We have trained this regression on 50% of the total number of towns in West Bengal and the remaining 50% is the test set. The Ridge coefficient i.e. alpha is taken to be 1.0.

IV. RESULTS AND DISCUSSION

The CNN showed am accuracy of 97.3% in classifying the daytime satellite images into high, medium or low intensities of nightlight intensity. The ridge regression shows some difference in accuracy for the two indices calculated in 3.2. The ridge regression model has an accuracy of 82 % when the Asset Index is calculated from the DCHB data.

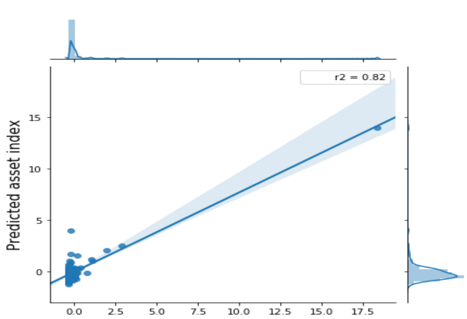


Figure 4. Plot showing relation of observed and predicted asset index from DCHB data

However, the ridge regression model has an accuracy of almost 93 % when the Asset Index is calculated from the Google Nearby Search results. This model is strongly predictive of asset index of any place provided its daytime satellite images are given as input into the model.

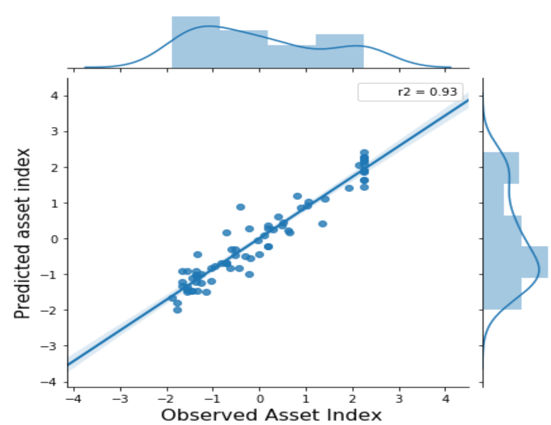


Figure 5. Plot showing relation of observed and predicted asset index from Google Nearby Search data

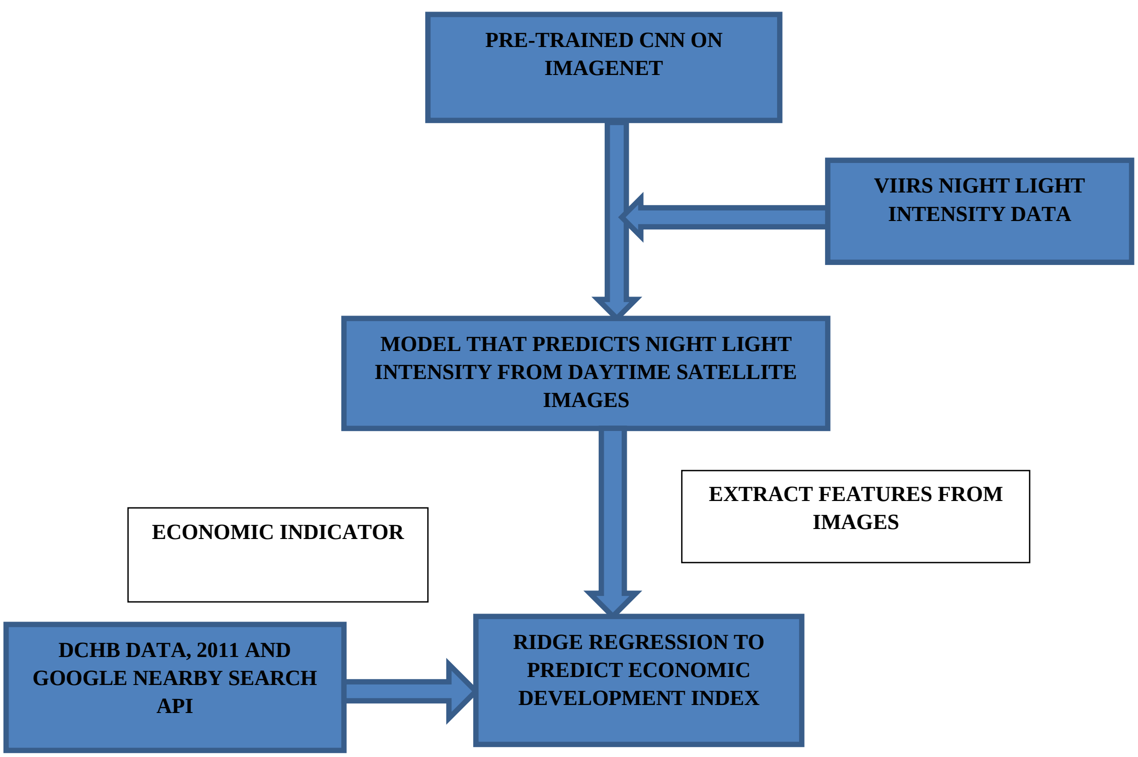
This shows that our predicted values are as very close to the observed ones.

As observed from the tables 3 and 5 the top few towns with highest asset index has changed in the two time periods. The two asset indices were normalized so that the values are between 0 and 1. Table 6 shows this data. We observe that Kolkata has retained its status and has the town with highest index number from the data of these two time periods. Siliguri, Kalimpong and Baranagar (M), Tamluk, Tarakeswar (M), Kanchrapara (M + OG), Durgapur are nowhere to be found in the top 10 from 2018 data compared to 2011 data. However places like Asansol, Jaypur(CT), Purulia, Barrackpore (M) and some other are new in the list.

TABLE 6: COMPARING THE TOP 10 TOWNS BASED ON ASSET INDICES.

List of Town Names for 2011	Normalized Asset Index_2011	List of Town Names for 2018	Normalized Asset Index_2018
Kolkata	1	Kolkata	1
Siliguri	0.171905278	Jaypur (CT)	1
Durgapur	0.123336303	Barasat	0.983679664
Haldia (M)	0.074732161	Bardhaman	0.675389524
Tarakeswar (M)	0.071583365	Purulia	0.660506183
Bardhaman	0.060212193	Asansol	0.629302507
Tamluk (M)	0.038969089	Baranagar (M)	0.565458158
Kanchrapara (M + OG)	0.029856469	Bally (CT)	0.564021163
Kalimpong	0.028314889	Chandannagar (M Corp)	0.547700827
Baranagar (M)	0.025793997	Barrackpore (M)	0.531380491

Figure 6. DATA MODEL



Our model trained to predict night light intensity from daytime satellite imagery outperforms that of Jean et al. Our R^2 value for the ridge regression is 82% and 93 % in the two cases for the towns of West Bengal compared to Jean et al. with model accuracy of 70% for 5 countries of South Africa. Our model definitely performs better. This might be due to availability of reliable structured data in India in contrary to that in South Africa. Kolkata is the largest town of West Bengal both in terms of population and financial activity. The largest 10 towns of West Bengal as cited in Wikipedia [8] include all the places as predicted our model. Hence the results from both sources of data are able to predict the asset index of these places very well.

V. CONCLUSION AND FUTURE SCOPE

From the accuracy of the ridge regression models trained on Census data and Google Nearby Search api data, we can conclude that we may use both the models to predict the asset index of the towns. Though this method is not 100 % accurate but still we can form a very good idea about economic activity of towns. The accuracy of the ridge regression model based on Google Nearby Search api data outperforms that of Census data. This shows that the former proves to be a good alternative source of data However we acknowledge the contribution of the Census data and it will always stay as the primary source of data for any such research in this field. We just presented an alternative way of calculating the Asset Index in the years when Census data is not available.

Further work can be done in this field by calculating such indices for every town of the country and comparing them over different time periods. This will show its development over the base year, which is in this case 2011. We have worked only on State; however this model can be implemented on any State of India as well as other countries. We chose town level data as we felt that towns are good representatives of economic activity. The towns can be categorized into a number of classes based on this asset index. Remote sensing technology is the future and a lot of social and geographical problems can be tackled with them.

ACKNOWLEDGEMENT

Tanuj Sur(TS), appreciate the help received from his parents Manidipa Pal Sur and Tarun Sur. Asoke Nath(AN) is grateful to Rev. Dr. Dominic Savio, Principal of St. Xavier's College for constantly motivating and supporting for doing research work in the field of Science and Engineering.

REFERENCES

[1] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, Trevor Darrell. DeCAF: "*A Deep Convolutional Activation Feature for Generic Visual Recognition*", 2013

[2] Neal Jean , Marshall Burke, Michael Xie, W. Matthew Davis, David B.Lobell, Stefano Ermon.*" Combining satellite imagery and machine learning to predict poverty*"

[3] Andrew Head, Mélanie Manguin, Nhat Tran, and Joshua E. Blumenstock. "*Can Human Development be Measured with Satellite Imagery?"* In Proceedings of ICTD '17, Lahore, Pakistan 11 pages, November 16–19, 2017.

[4]Abdisalan M Noor, Victor A Alegana, Peter W Gething, Andrew J Tatem and Robert W Snow. "*Using remotely sensed night-time light as a proxy for poverty in Africa"*, 2008.

[5] Nitish Shirish Keskar, Dheevatsa Mudigere, Jorge Nocedal, Mikhail Smelyanskiy, Ping Tak Peter Tang "*On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima"*,2017

[6]Map Box api code source: Jimut Bahan Pal https://github.com/Jimut123/apple_maps_api.

[7] Buddhadeb Ghosh, Prabir De "*Role of Infrastructure in Regional Development-A Study over the Plan Period",*1998.

[8]Major cities of West Bengal https://en.wikipedia.org/wiki/Cities_and_towns_in_West_Bengal#Major_cities

[9]District Census Handbook data,2011 http://censusindia.gov.in/2011census/dchb/DCHB.html

[10]Google Nearby Search api documentation https://developers.google.com/places/web-service/search#PlaceSearchRequests

Authors Profile

Tanuj Sur is a student of Dept. of Statistics, St. Xavier's College(Autonomous), Kolkata-700016. He is an aspiring deep learning researcher. Deep learning and machine learning are his primary interests. He has his own startup Less Than More which is an apparel merchandising online store. He has applied machine learning on this online platform to build recommendation engines and chat bots to improve user interaction. He participates in Kaggle Competitions and usually ranks in top 30%.

Dr. Asoke Nath is working as Associate Professor in the Department of Computer Science, St. Xavier's College (Autonomous), Kolkata. He is engaged in research work in the field of Cryptography and Network Security, Steganography, Green Computing, Big data analytics, Li-Fi Technology, Mathematical modeling of Social Area Networks, MOOCs etc. He has published more than **242** research articles in different Journals and conference proceedings.

