

# Artificial Intelligence in e-Governance Applications

## Extensive use of Machine Learning & Deep Learning

Artificial Intelligence development frameworks play a crucial role in enabling faster delivery of models built on them. They help provide data driven insights. Here, Machine Learning & Deep Learning Case Study is being considered for Object Detection, which may find an application in any other area where infrastructure assets are being created by the government.

**SHARMISTHA DAS GUPTA**  
Sr. Technical Director  
sharmi@nic.in



**ANY GUPTA**  
Scientific Tech. Asst. - A  
any.gupta78@nic.in

Edited by  
**MOHAN DAS VISWAM**

Artificial Intelligence (AI) is the intelligence demonstrated by machines in which computer systems perform tasks as humans like speech recognition, decision-making, language translation etc. AI has the ability to automate repetitive tasks, make connections, see relationships and make predictions with reasonable levels of accuracy. One requirement of using AI Technologies is to read and learn about various algorithms before applying as there are many options to choose from and being an emerging technology, new options are evolving fast.

The focus of India's National Artificial Intelligence Strategy is not only on economic growth, but also on social inclusion. Niti Aayog, in its National Strategy for Artificial Intelligence, has identified "AI for All" as the theme for leveraging full potential of AI to meet the country's unique needs and aspirations. It has identified five critical sectors for AI intervention, namely Healthcare, Agriculture, Education, Smart Cities & Infrastructure and Smart Mobility & Transportation.

Given the context of societal inclusion, in one of the key sectors of development i.e., Infrastructure creation for the citizens by ministries, NIC has extended support in Mission Mode Projects for faster delivery of services to citizens.

Citizens participated through web, Mobile Apps and Citizen Service Centres. It was decided to explore potential insights that can be gained in the effective implementa-

tion of these services, utilization of funds and creation of assets by using AI & Deep Learning.

The applicants upload photos of progress of work to get instalments from these portals through Direct Benefit Transfer. For establishing right utilisation of funds under these schemes, physical verification checks are conducted by officials before transfer of funds to a beneficiary's bank account. Sometimes, this leads to delay in timely disbursement of instalments. The subsequent part of the article will show how this cycle time is cut using AI modelling, and there will be a walkthrough of the process of AI model building.

### Technology Brief

#### Data readiness for use of AI in eGovernance

Next step was to assess data readiness of the organization for applying AI Technologies. Answers were required for the following questions:

- Is there sufficient data available for machine to learn patterns from?
- Is data being captured at source?
- Is it in a form that can be used as is or needs to be cleaned, scaled, or transformed?
- Is sufficient amount of annotated data available for supervised learning?

Swachh Bharat Urban was capturing geo-tagged constructed toilet photo images with help of a Mobile App developed by NIC. This provided for uniform 65K sized images that were being stored in a file folder, in addition to a copy in archived database. These files in folders were easily transferred for both Machine Learning Exercise and Deep Learning Model Building.

#### Platform availability for undertaking the exercise

It is a known fact that image processing is a compute intensive exercise and Deep Learning is the way to go. Centre of Excellence in Artificial Intelligence created the AI platform by acquiring two numbers of Deep Learning Server for AI Research (DGX I) with two Central Processing Units (CPUs) & eight Graphical Processing Units (GPUs) each and put them on

cluster for workload management with Kubernetes. Docker containers were installed for running Tensorflow, OpenCV for Deep Learning and TensorRT for inferencing.

### Annotated data availability

For the model to be trained using supervised learning algorithms, it is required to prepare a set of annotated training samples, randomly selected from the universe of unannotated image set. AI and Machine Learning have a limited ability to analyze data without labels. Same is true in context of image processing also. Hence, the images were annotated with bounding boxes both for Machine Learning & Deep Learning. Deep Learning requires a lot of annotated data. However, since there were only two classes i.e., detection of beneficiary and detection of toilet seats, few batches of 2000 - 3000 records each, both for Urban & Grameen toilets images could be trained with. Synthetic data was also created for hardening the model.

### How it worked?

#### Hypothesis for training model

AI modeling usually involves building a hypothesis based on features or attributes, which may be considered as independent variables and the hypothesis may be compared with the ground truth or the value of the dependent variable. There is an attempt to find a best fit between the model and the ground truth such that the residual error between the two is minimized and the hypothesis is an accurate predictor of the dependant variable.

The hypothesis in this model building

exercise is the classification of objects detected in an image as belonging to either of the two classes: beneficiary or toilet seat with at least 90% average precision for both the classes and a recall >0.95 (recall means if an object is there, it will be detected) to be identified as a good object detection model for the e-Governance application under consideration.

### Choosing the right algorithm

Model building is a heuristic exercise that often involves many rounds of iterations with different algorithms to see what gives the best insight from the data. Metrics are there to compare the results from these models. For Machine Learning exercise, HAAR Cascade Frontal Face Default Classifier was chosen as it can detect faces and has a very high detection (true positive rate) and very low false positive rate always. k-Nearest Neighbor (KNN) & Scale Invariant Feature Transform (SIFT) were chosen for prototyping in toilet seat detection as SIFT can robustly identify objects even among clutter and under partial occlusion. This is because the SIFT feature descriptor is invariant to uniform scaling, orientation and illumination changes, and it is partially invariant to affine distortion.

Deep Learning was tried with You Only Look Once (YOLOv3) using Darknet feature extractor. YOLOv3 is an open source Convolutional Neural Network (CNN) architecture framework, which uses a variant of Darknet. It is a 106 layer neural network trained on Imagenet. YOLO is a popular framework as it has

both accuracy and speed. Stochastic gradient descent was used to reach the optimum.

### Features

#### Model Building Exercise

Deep Learning model building for Swachh Bharat Urban was done using Transfer Learning with weights loaded from darknet53.conv.74, using 10% of training dataset as cross validation set randomly. The model performance was measured against this cross validation set also called test set. Similar exercise was done for Swachh Bharat Grameen for two rounds of iteration of pure model building. Then the parametric weights obtained for the model were used for hybrid model training of using Grameen weight on urban dataset and viceversa. Around 92% average precision was obtained through this exercise. Nearly 11000 images were annotated.

Then 5000 more images were created synthetically by cutting the toilet seat and beneficiary photos from original images and using scripting to put them randomly on different background. This put them in different positions, illuminating conditions and angles using translation, rotation etc. Hence, the model can be hardened for predicting data even in the absence of annotated dataset that modelled for these conditions. 100% mean Average Precision (mAP) was obtained for the model when weights from this annealed model on the actual dataset were used.

Here are the first and last couple of batches of model training.



Beneficiary and toilet seat detection in Swachh Bharat Scheme

Training dataset - Urban 1st batch	Training with starting weights darknet53.conv. 74	Beneficiary detection Average Precision percentage	Toilet seat detection Average Precision percentage	Remarks
1866 records	2000	75.84	83.27	It can be seen that object detection average precision oscillates for a few epochs and is best at 12000 weight. Then it starts to peter out. The reason is YOLOv3 uses mini batch Stochastic Gradient Descent to converge to local optima with slow learning rate.
	3000	73.83	81.96	
	8000	75.59	82.91	
	10000	76.83	85.53	
	11008	76.31	83.92	
	12000	77.25	85.23	
	13008	76.22	84.99	
	14000	76.19	85.89	
15008	75.69	85.79		

Table 1: Object detection by Deep Learning in Swachh Bharat Urban

## Technology Update

### Case I – Swachh Bharat Urban – Pure Model Training

Initial epochs were run on 1 GPU and then on multiple GPUs on DGX servers.

#### Note on Confusion Metrics

TP = True Positives, TN = True Negatives, FP = False Positive, FN = False

	Class 1 Predicted	Class 2 Predicted
Class 1 Actual	TP	TN
Class 2 Actual	FP	TN

Table 2: Predicted Classes

Negative, P - Precision, R- Recall, F1-Score means low false positives & low false negatives, IoU – Intersection over Union, mAP = mean Average Precision.

Since batch-wise stochastic gradient descent uses randomness and reaches local optima, after nine batches of training and cross-validating, annealing using synthetic data was used for hardening the process of model building. Then it was finally used on real dataset to reach global optima.

### Case II – Swachh Bharat Grameen – Hybrid Model Training with Annealing

Training dataset – Grameen 3rd batch	Starting with Urban-Grameen Synthetic Model weights yolo-obj_66024	Beneficiary detection Average Precision percentage	Toilet seat detection Average Precision percentage	Remarks
3000 records	67048	90.91	90.12	Cross validation is used by randomly shuffling between training and test set to increase performance of the model as seen next.
	68072	90.91	90.53	
	69000	90.91	90.04	
	70024	90.91	90.33	
	<b>72072</b>	<b>90.91</b>	<b>90.67</b>	
After Randomly shuffling 3000 records & using yolo-obj_72072	73000	100	100	P = 0.99, R = 1, F1-score = 1, TP = 707, FP = 5, FN = 0, IoU = 85.48 & mAP = 100%
	74024	<b>100</b>	<b>100</b>	<b>P = 1, R = 1, F1-score = 1, TP = 707, FP = 0, FN = 0, IoU = 86.79 &amp; mAP = 100%</b>
	75048	100	100	P = 1, R = 1, F1-score = 1, TP = 707, FP = 1, FN = 0, IoU = 86.46 & mAP = 100%

Table 3: Object detection, Grameen Hybrid Model Training

### Case III – Swachh Bharat Urban – Hybrid model training with Annealing

Training dataset – Urban 2nd batch	Starting with Urban-Grameen Synthetic Model weights yolo-obj_66024	Beneficiary detection Average Precision percentage	Toilet seat detection Average Precision percentage	Remarks
2075 records	67048	90.86	88.43	
	68072	90.91	88.87	
	<b>69000</b>	<b>90.91</b>	<b>89.45</b>	
	70024	90.91	81.42	
After randomly shuffling 2075 records & using yolo-obj_69000	70024	100	99.97	P = 1, R = 1, F1-score = 1, TP = 613, FP = 2, FN = 0, IoU = 82.83 & mAP = 100%
	71048	100	100	
	72072	100	99.97	<b>P = 1, R = 1, F1-score = 1, TP = 613, FP = 1, FN = 0, IoU = 84.29 &amp; mAP = 100%</b>
	73000	100	99.97	
	<b>74024</b>	<b>100</b>	<b>100</b>	
	75048	100	100	P = 1, R = 1, F1-score = 1, TP = 613, FP = 1, FN = 0, IoU = 83.56 & mAP = 100%

Table 4: Object detection, Urban Hybrid Model Training

## Impact

It will be a great help for the beneficiaries of government schemes if the system can prompt them to submit proper images so that chances of rejection of uploaded images get minimized. This will reduce pendency of applications for clearance, and consequently lead to speedier transfer of funds to citizens.

## Application Areas

A Mobile App has been developed to facilitate citizens to check the appropriateness of constructed toilet images uploaded. Using this AI model inference, the App can be used to check for beneficiary and toilet seat at the time of uploading of constructed toilet images and promptly alert users of improper photo uploads. Further experiment is being done on the face verification of applicant's photo with the beneficiary image in constructed toilet photo uploaded using FaceNet.

## Summary

The model built showcased how the building of such models can be taken ahead as it will be useful in detecting constructed and under construction government assets in similar schemes like in Prime Minister Awas Yojana, Atal Mission for Rejuvenation & Urban Transformation etc., by training in a similar manner. ■

For further information, please contact:

**SHARMISTHA DAS GUPTA**  
Senior Technical Director  
Centre of Excellence in AI, A3B8 Bay  
NIC HQ, A-Block, CGO Complex  
NEW DELHI - 110003

Email: sharmi@nic.in  
Phone: 011-24305211