

# Details of the Submission Format (Dhaka-AI 2020)

## Instruction to generate the CSV file

### Key attributes

For this competition, there are 21 classification labels over a diverse set of images. The key elements for evaluations metrics are:

- Classification Labels
- Confidence
- Bounding Boxes

### Key metrics

This competition is evaluated on the mean average precision at a different intersection over union (IoU) thresholds. The IoU of a set of predicted bounding boxes and ground truth bounding boxes is calculated as:

$$IoU(A, B) = \frac{A \cap B}{A \cup B}$$

The metric sweeps over a range of IoU thresholds, at each point calculating an average precision value. The threshold values range from 0.5 to 0.75 with a step size of 0.05. In other words, at a threshold of 0.5, a predicted object is considered a "hit" if its intersection over union with a ground truth object is greater than 0.5.

At each threshold value  $t$ , a precision value is calculated based on the number of true positives (TP), false negatives (FN), and false positives (FP) resulting from comparing the predicted object to all ground-truth objects:

$$\text{precision}(t) = \frac{TP(t)}{TP(t) + FP(t) + FN(t)}$$

A true positive is counted when a single predicted object matches a ground truth object with an IoU above the threshold. A false positive indicates a predicted object had no associated ground truth object. A false negative indicates a ground truth object had no associated predicted object.

The average precision of a **single classification label** is calculated as the mean of the above precision values at each IoU threshold:

$$\frac{1}{|IoU\ Thresholds|} \sum_t Precision(t)$$

In your submission, you are asked to provide classification labels for each bounding box as well as a confidence score. At first, the class labels are used to match the target labels then bounding boxes will be evaluated based on the above process. This means that bounding boxes with correct labels

will be checked with target labels and bounding boxes, which determines what boxes are considered true positives.

## Submission File

The submission format requires a comma delimited set of bounding boxes. For example:

```
image_id_01,Truck,0.5,766,230,904,552,1024,1024
```

indicates that image `image_id_01` has a bounding box with a confidence of 0.5, at `xmin == 766`, `ymin == 230`, `xmax == 904`, `ymax == 552`, with width and height of 1024 by 1024.

It is important to scale the images to 1024x1024 as all the test images are scaled to 1024.

The file should contain a header and have the following format. Bounding boxes are formatted using pascal voc.

	image_id	class	score	xmin	ymin	xmax	ymax	width	height
	Numan_(101)_jpg.rf.a9abd780ae62e7aa8a861c7693d...	rickshaw	0.79	766	230	904	552	1024	1024
	Numan_(133)_jpg.rf.1e1ec524e09e67f50afdc03fd3a...	rickshaw	0.93	535	110	648	435	1024	1024
	Numan_(133)_jpg.rf.1e1ec524e09e67f50afdc03fd3a...	bus	0.73	467	28	594	200	1024	1024
	Numan_(133)_jpg.rf.1e1ec524e09e67f50afdc03fd3a...	pickup	0.50	647	99	723	205	1024	1024
	Numan_(133)_jpg.rf.1e1ec524e09e67f50afdc03fd3a...	rickshaw	0.93	535	110	648	435	1024	1024

## Sample Code

[https://colab.research.google.com/drive/1x2NZUVJtVqmeSFpBO\\_XNmV2esMrHPGQ7](https://colab.research.google.com/drive/1x2NZUVJtVqmeSFpBO_XNmV2esMrHPGQ7)

Here is a sample solution provided for a quick start. The submission format is described in detail on a separate page. Also from the sample solution, you may check out the code snippet for the submission file structure.

