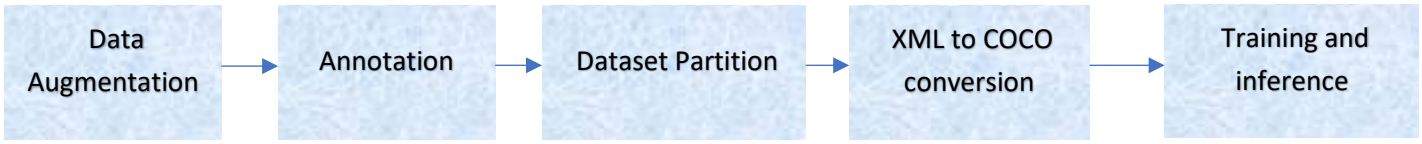
OBJECT DETECTION

The objective is to develop a deep learning algorithm to detect objects in an image/video frame and extract the bounding box(es).

I have used Detectron2 platform for object detection, introduced by Facebook AI, and also uses Faster R-CNN algorithm in the background. I have taken 25 PAN card images from the internet.



1. PAN card images are downloaded from the internet (25 images) and are augmented by added grayscale and blur transformations, totaling to 75 images.
2. The augmented dataset is annotated using makesense.AI platform with bounding boxes drawn over Permanent Account Numbers on the image, with 'PAN' as the label. (Script: data_aug.py)
3. The annotations are exported as XMLs' and dataset has been partitioned with 70% for training, 15% for validation, and the remaining 15% for testing. (Script: partition_data.py)
4. The XMLs' from train, validation and test datasets are converted into COCO format to store the annotations in JSON, describing object classes and bounding boxes. (Script: voc2coco.py)
5. The partitioned dataset (both images and COCO format JSON) is imputed into the training pipeline. The pipeline uses Detectron2 platform with Faster R-CNN algorithm running in the backend. (Notebook: PAN_Card_Detectron_Training_Inference.ipynb)
6. Basic hyperparameters:
   - ➔ Iterations: 1500
   - ➔ Learning Rate: 0.001
   - ➔ Batch Size: 4
   - ➔ Number of classes: 1
7. Model is automatically saved after training and is utilized for predictions in the inference. The bounding boxes are also extracted as a part of the objective.