**AWS Machine Learning Engineer Nanodegree Capstone Project**

By Araz Sharma 22nd February 2022

1. **DEFINITION**

**Project Overview:**

# This Project is on “Inventory Monitoring at Distribution Centres”. This problem has originated from the popular field of Robotics, and its applications in the industry, like that of Inventory Monitoring and Supply Chain Functions. A lot of Corporations, which handle physical cargo and deal with supply chain of any kind of goods have tried to bring in automation to make these processes more efficient and accurate. A great example of this is Amazon, who is one of the biggest hubs of delivery of all kinds of goods. These goods are often stored in big warehouses. Since the quantity of these items are in a huge amount, to physically do inventory monitoring would require both large and intelligent human resources, which are both expensive and prone to errors.

# This is where robots come in to help in Inventory Monitoring. They can be trained with Machine Learning Models, to perform tasks like Object Detection, Outlier & Anomaly Detection and much more. Once trained, these models are scalable, and can be deployed at a low cost for usage in actual warehouses and distribution centres on industry level robots.

**Problem Statement:**

# As mentioned above in the domain, distribution centres often have robots which carry objects. These objects are present in bins, and for our problem, each bin can contain 1-5 objects.

# The problem we aim to tackle in this project is to count the number of items present in the bin. This is a Classification task, of classifying number of items in 1 – 5, given an input image. This is a worthwhile problem to solve, for it has immense real world applications. If we can develop a model, which can take in a picture of a bin, and accurately return the number of objects present in that, we could solve & thus fully automate one crucial step in the Inventory Management process!

# Datasets and Input:

# The dataset used in this problem is the open source Amazon Bin Image Dataset. This dataset has 500,000 images of bins containing one or more objects present in it. Corresponding to each image, is a metadata file, which contains information about the image, like the number of objects it has, the dimensions and type of objects. For our problem statement, we only need the total count of objects in the image. An example of an image & the corresponding metadata file is shown as below: (Source [Reference [1]](https://github.com/awslabs/open-data-docs/tree/main/docs/aft-vbi-pds))

# The “EXPECTED\_QUANTITY” field tells us the total number of objects in image.

# A picture containing text, indoor Description automatically generatedTable Description automatically generated

# Solution Strategy:

# To solve our problem statement, we will use the task of Computer Vision, to come up with a Machine Learning Model, which given an image from our dataset, can identify the number of objects present in it. Essentially, we would be using Multi-Class Image Classification, with Number of Objects from 1-5 as an individual class.

# To do this, we can make use of Convolutional Neural Networks, which are a State of the Art technique for Image Recognition tasks. We will leverage Pre-trained models with Transfer Learning to solve the problem.

# The end solution should be a model, which can take in an input image from the Amazon Bin Image dataset, and accurately output the number of objects it thinks are present in that image (from 1-5).

# Metrics for Evaluation:

# To evaluate our model, we need some good metrics, which align with the problem statement. The metrics must be mathematically sound and we should be able to optimise our model for them.

# Since we have a Classification Task, we can use Accuracy, Recall, Precision and F1 scores as our metrics. These can be for the overall data, and also class-wise, to identify if a model is doing better on a particular class, or has a high bias for one of them.

# Since we have Multi-Class classification, the definition of Precision & Recall must be clearly understood. Let’s assume we have 3 classes: A, B & C.

# With Multi-Class classification, we have precision and recall for each individual class. For example, metrics for Class A will be defined as follows:

# Precision: (‘#’ stands for Number of)

# #Correctly Predicted Class A Instance / #All Instance predicted as Class A

# Recall: (‘#’ stands for Number of)

# #Correctly Predicted Class A Instances / #Total Class A Instances

# F1 Score: 2\*Precision\*Recall / (Precision + Recall)

# So to evaluate our model, we will see per class metrics of Precision, Recall and F1 Score, and see the overall accuracy of the model, which includes all classes, and is defined as:

# Accuracy: #Total Correctly Predicted Instances / #Total Instances

# [Reference [3]](https://towardsdatascience.com/multi-class-metrics-made-simple-part-i-precision-and-recall-9250280bddc2) for Precision, Recall and F1 for multi-class classification