

# Machine Learning Engineer Nanodegree

---

## Capstone Proposal

*Khang Hoang Tran July 2nd, 2024*

## Proposal

### Domain Background

In the logistics domain, efficient inventory management and accurate order fulfillment are critical for maintaining operational efficiency and customer satisfaction. Many activities in warehouse these days involve much human resource. For large-scaling business, instantly increasing labor force might produce a greater cost with a potential risk of staff management. As a result, the rapid growth of e-commerce and the increasing complexity of supply chains have heightened the need for advanced technological solutions to streamline these processes. One such solution has been popularly adopted is leveraging technologies like machine learning techniques to automate and enhance inventory monitoring, recognizing and sorting tasks.

### Problem Statement

In Amazon warehouse scenario, robots are used to move objects as a part of their operations. Objects are carried in bins which can contain multiple objects. For each customer's order, we need to ensure that delivery consignments have the correct number of items. Amazon robotics, scanning machines, and computer systems in fulfillment centres can track millions of items in a day. A tracking inventory system with only manual checks are labor-intensive.

This project focuses on utilizing the "Amazon Bin Image Dataset" to develop a machine learning model capable of recognizing and specifying the number of objects in each bin. The dataset comprises a diverse collection of images captured in a warehouse setting, depicting various items placed in bins. By applying advanced image recognition and classification algorithms, the goal is to create a robust system that can accurately identify products, manage inventory levels, and optimize order picking processes. Also, the output model aims to reduce manual labor, minimize errors, increase overall efficiency; and most important, simulate a full machine learning pipeline in a logistic data-processing job.

### Datasets and Inputs

Official dataset reference: [Amazon Bin Image Dataset](#)

The Amazon Bin Image Dataset contains over 500,000 images and metadata from bins of a pod in an operating Amazon Fulfillment Center. The bin images in this dataset are captured as robot units carry pods as part of normal Amazon Fulfillment Center operations.

As for a large dataset, we plan to use only a subset of **XXX** samples from the original dataset which is served as experiment tasks to examine the efficiency of the model. The subset only contains bin images that store the number of items between '1' and '5'.

Sample of the images.



The subnet data have 5 subfolders corresponding to 5 labels, each subfolder contains images. In regard to build the model, we divide our dataset into train, test and validation set. Our final dataset structure looks like this

```

+---bin-images
  +---train
    |   +---1
    |   |       00014.jpg
    |   |       00024.jpg
    |   |       ...
    |   +---2
    |   |       00056.jpg
    |   |       00112.jpg
    |   |       ...
    #   ...
    |   +---5
    |       00006.jpg
    |       00058.jpg
  +---test
  ... (same as above)
  +---validation
  ... (same as above)

```

For each image there is a metadata file containing information about the image like the number of objects, it's dimension and the type of object. An image, i.e `00014.jpg` will have metadata file `00014.json`.

Sample of the metadata

```

{
  "BIN_FCSKU_DATA": {
    "B0123H9HME": {
      "asin": "B0123H9HME",
      "height": {
        "unit": "IN",

```

```

        "value": 3.49999999643
      },
      "length": {
        "unit": "IN",
        "value": 9.599999990208
      },
      "name": "Organix Grain Free Lamb & Peas Recipe, 4 lb",
      "normalizedName": "Organix Grain Free Lamb & Peas Recipe, 4 lb",
      "quantity": 1,
      "weight": {
        "unit": "pounds",
        "value": 4.549999996184413
      },
      "width": {
        "unit": "IN",
        "value": 8.099999991737999
      }
    }
  },
  "EXPECTED_QUANTITY": 1
}

```

We have got one of the problems dealing with this dataset. Some small boxes are inside other larger boxes which can disrupt our predictions. So for an object recognition task, we will need to clean the dataset. Since we already have `file_list.json`, we are allowed to proceed further without spending time with initial data. So the metadata is just for references or if we want to collect more samples for our dataset.

## Solution Statement

From introduction's description, our problem is related to image classification and object recognition; therefore, the relevant model would be utilized to support us to reach the target. The solution will comprise of multiple stage, but the core model is a pre-trained(Resnet 50) deep learning model. The model will be trained with the sufficient images from the mentioned dataset above and for training phase, we use *train* and *validation* set. The input data that is fed to the model is the raw content of an bin image which contains a number of objects. The output should be a *list* of the predicted score(probabilities) for all 5 labels(number) of objects in each bin.

We will use appropriate AWS cloud services to create a complete machine learning pipeline: storing data, training model, hosting instance for inference, ... For the model learning tasks, we take advantage of SageMaker as an experiment environment which is convenient for us to mainly concentrate on improve model and workflow.

The model will be evaluated using **accuracy** as a metric. The trained model will be tested against a *test* set to observe objectively the performance and the precision of the final result.

## Benchmark Model

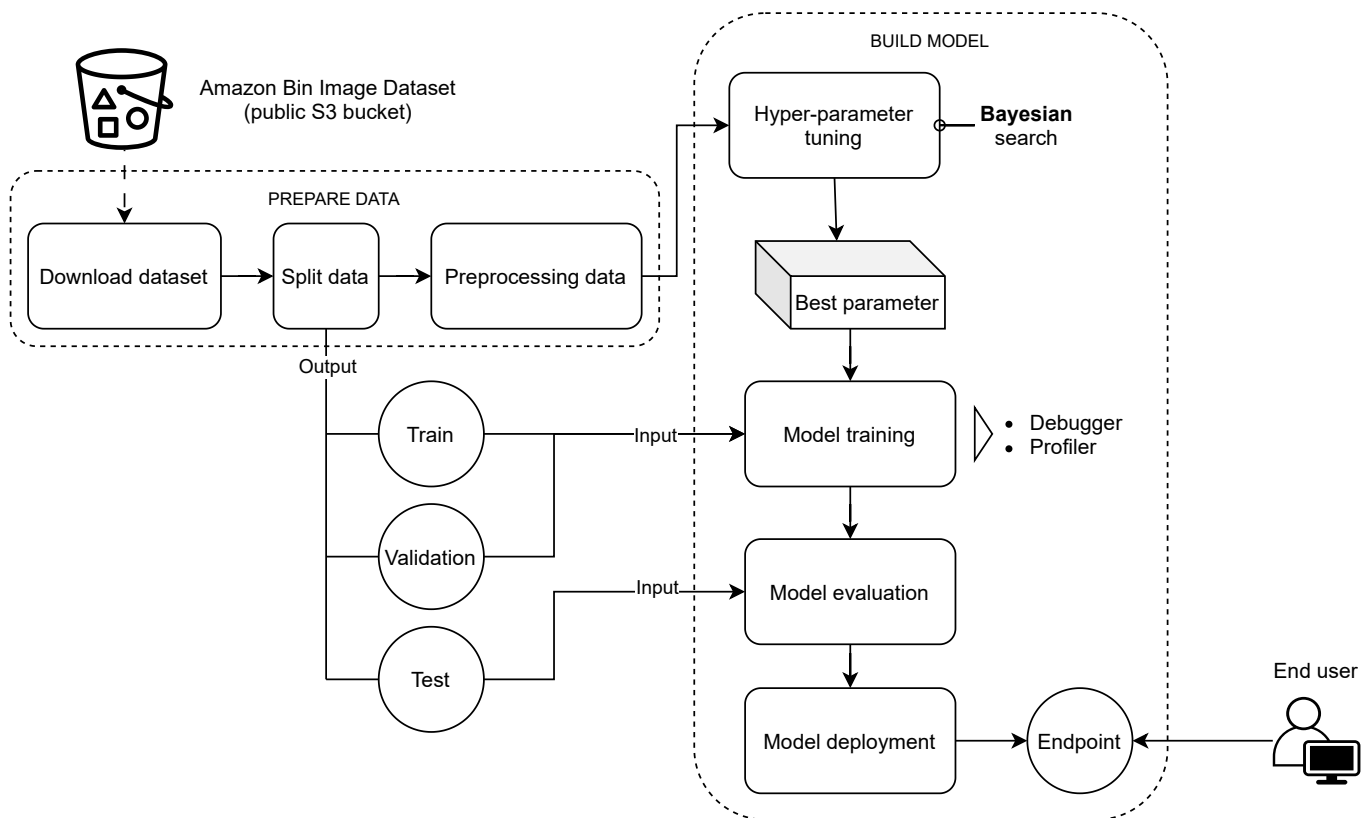
The project will be benchmarked using accuracy as we determined since it is a classification problem. Our target is to achieve an accuracy of above **55.67%** on validation set **using Deep Learning models** which is the final and accepted value in the [Amazon Bin image dataset challenge](#) by [silverbottlep](#)

## Evaluation Metrics

For a classification problem, we usually assess the level of correctness of model by **accuracy** as a metric. Moreover, there are other metrics that enable to supervise model's accuracy and detect abnormal point in the dataset or final result. For that reason, we could use the **Confusion matrix** and common metrics such as precision, recall, F1-score, ROC/AUC curve.

## Project Design

An overview of process for building the system:



Steps are performed as following:

1. Download a subset of dataset from public S3 bucket with **file\_list.json**
2. Divide data for train, test and validation, then upload these dataset to specific S3 bucket for this project.
3. Preprocessing data (if needed) before data is fed to model as valid input
4. Apply hyperparameter optimization strategy to find the best parameter set.
5. Use the best combination of parameters above as a input to train the final core model
6. Keep track, monitor and continuously evaluate model with test set. If the metric's values produce a optimistic result, then go to next step.
7. Deploy model as an endpoint, create Lambda function for application to use for inference

## Reference

- Dataset: [Amazon Bin Image Dataset](#)
- Dataset Usage 1: [Amazon Inventory Reconciliation using AI](#) by [pablo-tech](#)
- Dataset Usage 2: [Amazon Bin Image Dataset\(ABID\) Challenge](#) by [silverbottlep](#)
- [Metrics to evaluate classification models](#)