AWS Machine Learning Engineer Nanodegree Program

Name: Luiz Otávio
Project: Inventory Monitoring at Distribution Centers

A project proposal encompasses seven key points:

- The project's **domain background** — the field of research where the project is derived;
- A **problem statement** — a problem being investigated for which a solution will be defined;
- The **datasets and inputs** — data or inputs being used for the problem;
- A **solution statement** — the solution proposed for the problem given;
- A **benchmark model** — some simple or historical model or result to compare the defined solution to;
- A set of **evaluation metrics** — functional representations for how the solution can be measured;
- An outline of the **project design** — how the solution will be developed and results obtained.

Capstone Proposal

Domain Background

Inventory monitoring is crucial and important in a distribution centers. The way you manage and monitor inventory in a distribution center influences how much you spend on logistics costs and how efficient you can get. [1] This process makes sure that each bins carry the right number of objects.

Tracking things manually was the norm before the industrial age. The earliest form of inventory management dates back over 50,000 years in which people used “tally sticks” to count. [2]

It is now possible to improve customer satisfaction, prevent dead stock, monitor and optimize stock levels in inventory management by implementing machine learning.

Problem Statement

There are a lot of labor requirements associated with manual inventory management, as well as **the** possibility of errors. So, there is a need for an effective inventory management.

Proper warehouse inventory management can save you money, including labor, storage, and fulfillment costs. This can be done by optimizing storage, investing in warehouse automation, organizing inventory to speed up the picking and packing process, and implementing technology to automate tasks that reduce human error. [3]

Machine learning is one technology that can automate those processes. Computer vision can be used to automate inventory monitoring using convolution neural network to count the number of objects in each bin and ensure that delivery consignments contain the correct number of items.

Datasets and Inputs

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. [4]

Since this is a large dataset, a sample of the dataset is used. The dataset is available on Amazon S3.

There will be subfolders for the data subset created. The number of objects in each of these subfolders equals the name of the folder. In folder 1, for example, there is exactly one object in every image. Afterwards, training, testing, and validation datasets will be developed.

Example of image:

Example metadata (json):


```
{
 "BIN_FCSKU_DATA": {
  "B00CFQWRPS": {
   "asin": "B00CFQWRPS",
   "height": {
    "unit": "IN",
    "value": 2.399999997552
   },
   "length": {
    "unit": "IN",
    "value": 8.199999991636
   },
   "name": "Fleet Saline Enema, 7.8 Ounce (Pack of 3)",
   "normalizedName": "(Pack of 3) Fleet Saline Enema, 7.8 Ounce",
   "quantity": 1,
   "weight": {
    "unit": "pounds",
    "value": 1.8999999999999997
   },
   "width": {
    "unit": "IN",
    "value": 7.199999992656
   }
  },
  "ZZXI0WUSIB": {
   "asin": "B00T0BUKW8",
   "height": {
    "unit": "IN",
    "value": 3.99999999592
   },
   "length": {
    "unit": "IN",
    "value": 7.899999991942001
   },
   "name": "Kirkland Signature Premium Chunk Chicken Breast Packed in Water, 12.5 Ounce, 6 Count",
   "normalizedName": "Kirkland Signature Premium Chunk Chicken Breast Packed in Water, 12.5 Ounce, 6 Count",
   "quantity": 1,
   "weight": {
    "unit": "pounds",
    "value": 5.7
   },
   "width": {
    "unit": "IN",
    "value": 6.49999999337
   }
  },
  "ZZXVVS669V": {
   "asin": "B00C3WXJHY",
   "height": {
    "unit": "IN",
    "value": 4.330708657
   },
   "length": {
    "unit": "IN",
    "value": 11.1417322721
   },
   "name": "Play-Doh Sweet Shoppe Ice Cream Sundae Cart Playset",
   "normalizedName": "Play-Doh Sweet Shoppe Ice Cream Sundae Cart Playset",
   "quantity": 1,
```

```
    "weight": {
      "unit": "pounds",
      "value": 1.4109440759087915
    },
    "width": {
      "unit": "IN",
      "value": 9.448818888
    }
   }
  },
 "EXPECTED_QUANTITY": 3
}
```


Solution Statement

To solve this image classification problem, it will be used computer vision techniques, in this case a pre-trained (Resnet 50) deep learning model. As input, we provide an image of a bin containing products and as output, we give predicted scores for each category type, for instance, number of objects. All data will be storage at S3.

The algorithm will be built using Convolution Neural Netwook (CNN) and the framework Pytorch. To improve the performance of the model, SageMaker's Hyperparameter Tuning will be used. After that, SageMaker endpoint will be created to deploy the model.

The SageMaker Studio will be used as our environment.

Benchmark Model

The dataset has two benchmark that can be found at Registry of Open Data on AWS. The publication Amazon Bin Image Dataset Challenge by silverbottlep archive an accuracy of 55.67%.

Evaluation Metrics

Since this is an image classification task in machine learning, accuracy is used to evaluate the model performance.

Project Design

The entire project will be built in the AWS environment. A pre-trained model will be used to classify the number of objects in each bin.

For this classification problem, the steps are:

- Create a new SageMaker Studio environment;
- Download the Amazon Bin Images Dataset and upload them to an S3 bucket. The data will be split into training, validation and testing sets.
- Create the script to train the model and do hyperparameter tuning. The model is a pre-trained model (Resnet50) using the framework Pytorch;
- Fine-tuning the pre-trained model;
- SageMaker's debugger and profiler will be used to measure log metrics and provide insights;
- Deploy the model using SageMaker endpoint.

References

[1] Lopienski Kristina, 2021. Warehouse Inventory Management Processes You Need to Know in 2022

[2] Writing Intern, 2018. HISTORY OF INVENTORY MANAGEMENT TECHNOLOGY

[3] Kristina Lopienski, 2021 Warehouse Inventory Management Processes You Need to Know in 2022

[4] amazon, Amazon Bin Image Dataset https://registry.opendata.aws/amazon-bin-imagery/