Warehouse Storage Optimization - Report 3

Dhruvil Dave, Dhatri Kapuriya, Harvish Jariwala, Nisarg Thoriya
School of Engineering and Applied Sciences
Ahmedabad University
Ahmedabad, Gujarat, India - 380009
dhruvil.d@ahduni.edu.in (AU1841003)
dhatri.k@ahduni.edu.in (AU1841129)
harvish.j@ahduni.edu.in (AU1841050)
nisarg.t@ahduni.edu.in (AU1841142)

***Abstract*—This is the third progress report of our group *Gopher - Group 4* for Machine Learning (CSE523) course project.**

***Index Terms*—Time series forecasting, classification, data pre-processing, clustering, graphical models**

I. Introduction

For our project, we decided to use the Amazon Bin Images Dataset.

This is originally a Computer Vision dataset. The Amazon Bin Image Dataset contains over 530,000 images and metadata from bins of a pod in an operating Amazon Fulfillment Center.

II. Brief insight of previous work

Our dataset has columns defining the features of products stored in the bins along with their respective locations in those bins. In the previous week, we performed exploratory data analysis on the dataset from which we inferred that volume of each bin will define the number of products each bin could hold. As we are not provided with the bin size, volume of products and their quantity would help us to decide maximum bin size possible in the warehouse and fractional portion occupied.

III. Task Performed And Outcomes

After a brainstorming session we landed to some important takeaways which are useful for future work. Our warehouse is in an empty state or lets say any state where some bins are filled partially or completely. Now, when any new product is to be added it would be dependent on the past data available i.e the statistics of empty or partly filled bins. On arrival of each product the model has to take the past state of the boxes into consideration along with the volume of the product. This resembles the time series forecasting where future events are predicted by analyzing the trends of the past, on the assumption that future trends will hold similar to historical trends.

Also, the model has to classify the incoming storage object into the bin it will be placed inside the warehouse. For that, we consider it to be a time series classification problem.

Let us take an array of length N to represent the bins. Here, N is the total number of bins available for us to store the objects in. We will consider that all the bins are uniform. So the maximum value of a bin will be fixed to us (something we found out from the data in the previous report). To represent the percentage of space remaining in the bin, we will use a relational matrix i.e. the number in the matrix will depict the fraction of empty bin.
Note: We use this fractional representation because this approach will help us maintain uniformity across different measurement system i.e. metric and imperial. This implies that when a fraction is one it would be an empty bin and when it would be zero, the bin must be filled completely.

The empty bin vector will look like this:

$$\text{Bin vector} = \{1, 1, 1, ...1\}$$

Now the first product arrives.

$$\begin{aligned}\text{Bin vector} &= \{1, 1, 1, ...1\}\text{+ Volume of first product} \\ &= \{1, 1, 0.8, 1, ...1\}\end{aligned}$$

IV. Tasks for Upcoming Week

The main tasks to be performed in the upcoming week are:

1) Time series forecasting
2) Modelling