Environmental Impact of Online Products

Marium Sarah   
*College of Business & Economics*  
*California State University, East Bay*Hayward, CA marium.sarah@csueastbay.edu  
  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID  
  
line 1: 4th Given Name Surname  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCI

*Abstract*—A solution to empower customers, to reduce the amount of plastic ending up in landfill and leading to an increase in greenhouse gas emissions, by accessing the environmental impact statistics and brand’s eco-friendly status of online products.

Keywords—environment, sustainability, image processing

# Introduction

Plastic is durable, cheap, and readily used to contain various products. Although plastic is accessible, it is also non-degradable and made to last over 400 years. A staggering 85% of the plastic thrown away in recycle bins end up in landfill, and the process of recycling the other 15% takes 16M tonnes of greenhouse gas emissions which is equivalent to 3.5M vehicles on the road [1]. The goal of this project is to alert customers of the environmental impact of products they purchase from online stores. The focus of this project is to estimate the level of plastic content in Amazon products and analyze the carbon footprint of the products and at-scale-businesses selling on Amazon.

## Project Overview

This project uses web-scraping of amazon products extracting detailed-description of weight and size, image processing to identify the type of plastic in products, statistical analysis and data visualization to compare and analyze the environmental impact of various brands.

## Problem Statement

Plastic is the most persistent non-degradable pollutant on Earth. The level of plastic packaging in products found online is rarely mentioned, thus estimating environmental impact is difficult. Knowing that majority of the plastic thrown away in recycle bins ends up in landfill, over 78% customers (surveyed by GreenPrint) hope to find eco-friendly products however they have a difficult time identifying them [2]. This project will help customers compare similar products based on their plastic content and the manufacturing brand’s environmental impact.

# Execution

Before we proceed to understanding the various data points to help estimate the carbon footprint of products. We will first produce an in-depth understanding of the project execution plan.

## Execution Plan

Fig. 1. highlights the inner workings all the way from the input to the final output of the program. The following are the 4 highlighted feature requirements of this project.

1. Web Scraping Amazon: Selected Product Categories will be scraped from Amazon to treat as input to our program.
2. Plastic Detection Model: Detect if the input Product Image contains Plastic or Not. Using an existing dataset for training and validation. Followed by testing directly on Amazon Product Images.
3. Plastic Level Algorithm: Algorithm to calculate the carbon footprint based on the volume and quantity provided in the Amazon Product Image Description.
4. An optional 4th existing model will be used to further identify if an image that is not Plastic is Recyclable or not. This will further help users understand that recycling also has a negative impact.
5. 1-3 Execution Requirements
6. Execution Plan Workflow

## Execution Steps

The following table highlights the execution steps that will be involved in the 12-week Capstone Project.

1. Execution Plan

| S.No. | Execution Plan | |
| --- | --- | --- |
| Step | Description |
| **1** | **Web Scraping** | Selection of certain Amazon Product Categories,  Web scraping over 1000 product details using beautiful soup package on Python. |
| **2** | **EDA & Data Cleaning** | Analyse the extracted data based on product type, popularity and any environmental impact mentions in description. Data cleaning to remove products without weight and size provided, and those without proper images provided. |
| **3** | **Partitioning & Pre-processing** | Various Pre-processing of Image data will be conducted to select the best data reduction methods. Including image cropping, edges detection, truncation. Furthermore, filters including Partition membership, discretization, normalization will be tested. Data partitioning will be used for the next step of predictive modelling. |
| **4** | **Predictive Modelling** | Create and compare common image detection predictive models including YOLO, CNN, RCCN. |
| **5** | **Compare Modelling Performance** | Cross-validation will be used along with evaluative metrics including accuracy, precision and recall comparing the model’s performance and choose the best model for our final prediction |
| **6** | **Visualize Product-level results** | Using statistical insights and common formulas of estimating environmental impact we will use various plots and graphs to create a standard easy to understand methodology of the eco-friendly level of any given product. |
| **7** | **Data Visualization Dashboard** | Using Tableau to visualize the environmental impact of various brands. This will include plots and graphs to highlight pattern and compare the performance of brands. |

## Strategy

Given the above execution plan, the strategy in place will involving performing Step 1 (Web Scraping) and Step 2 (Data Cleaning & EDA) in a Waterfall methodology one after the other in the first 2 weeks of the project.

Followed by utilizing a Sprint-based Agile Methodology focused on choosing the best Predictive Models (Steps 3-5), each sprint will start with creating the model, applying different combinations of filters and pre-processing techniques and evaluating the model’s performance against previous models tested.

Finally Step 6 and Step 7 compiling the Data Analytics and visualization will conclude the project completion. This will be completed in the final 2 weeks of the project.

# Data Description

The test data produced will be extracted from the Amazon website (referring to Fig. 1. 1st Requirement – Web Scraping Amazon) from specific categories focused on common household items that are purchased frequently. This dataset will have over 1000+ items with the following fields:

* **Product Name**
* **Image Link**– Link to image to extract pixels
* **Brand-** Brand manufacturing the given product.
* **Units-**Number of units within the product
* **Size-**The size ofitem
* **Item dimensions:** Given in inches LxWxH
* **Unit count:** The number of units/items in the given product
* **Item Weight:** The weight of the product
* **Item Volume:** For liquid items, the volume of the contentgiven in the packaging
* **About this item:** A detailed description about the specifications of the item on sale
* **No. of Ratings:** Total number of ratings given to this item
* **Ratings:** The overall rating on the item (scale of 1-5)
* **Price:** The price in dollars of the item

##### Data Collection

To develop a strong Image Detection Model (referring to Fig. 1. 2nd Requirement – Plastic Detection Model) we first need to research and identify datasets with a proper differentiation of Plastic and Non-Plastic products.

Most of the relevant Plastic detection datasets are focused on images from Waste Facilities and Litter. One challenge when picking appropriate training datasets is to consider mainly those that consist of unused new product images rather than Waste or Litter. The following table consists of researched relevant datasets.

1. Table Type Styles

| Dataset | Plastic and Non-Plastic Datasets | | |
| --- | --- | --- | --- |
| Images | Classification | Comments |
| Glassense-  Vision [11] | 500 | banknotes, cereals, medicines, cans, sauces, bottles, deodrants |  |
| Drinking Waste Classification [7] | 25.4k | Aluminium, cans, glass, bottles, Plastic bottles, Milk bottles |  |
| Waste Classification Data [5][6] | 26k | Organic (12.6k), Recyclable (10k),  Non-Recyclable (3k) |  |
| Trashnet [4] | 2k | plastic (482)/ trash, paper, metal, glass, cardboard (1.6k) |  |
| Trashbox [10] | 17k | Plastic, e-waste, medical waste, paper, metal, glass, cardboard |  |

##### Metrics and benchmark

All the developed models will firstly be applied on the test data. We will then compare the performance of the various models that are tested first based on the model accuracy, approving models with an accuracy over 80% and choosing the best model based on the precision, recall and F1 score.

##### References

1. Margaret Osborne, “At Least 85 Percent of U.S. Plastic Wate went o Landfills in 2021,” Smithsonian Magazine, May 9, 2022. *(*[*Link*](https://www.smithsonianmag.com/smart-news/the-us-recycled-just-5-percent-of-its-plastic-in-2021-180980052/)*)*
2. Michele Koch and Zach Harris, “GreenPrint Survey Finds Consumers Want to Buy Eco-Friendly Products, but Don’t Know How to Identify Them,” Business Wire, March 22, 2021 ([Link](https://www.businesswire.com/news/home/20210322005061/en/GreenPrint-Survey-Finds-Consumers-Want-to-Buy-Eco-Friendly-Products-but-Don%E2%80%99t-Know-How-to-Identify-Them))
3. <https://github.com/AgaMiko/waste-datasets-review>
4. <https://github.com/garythung/trashnet>
5. <https://www.kaggle.com/techsash/waste-classification-data> - ORGANIC vs RECYCLABLE
6. <https://www.kaggle.com/sapal6/waste-classification-data-v2> INCLUDES NON RECYCLABLE
7. <https://www.kaggle.com/datasets/arkadiyhacks/drinking-waste-classification> clear background
8. http://www.slipguru.unige.it/Data/glassense\_vision/
9. DATASET RESEARCH – ONE <https://github.com/antiplasti/Plastic-Detection-Model> GOOD BUT NOT REALLY ONLINE PRODUCTS
10. <https://github.com/nikhilvenkatkumsetty/TrashBox>
11. http://www.slipguru.unige.it/Data/glassense\_vision/
12. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
13. K. Elissa, “Title of paper if known,” unpublished.
14. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
15. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
16. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
17. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
18. K. Elissa, “Title of paper if known,” unpublished.
19. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
20. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
21. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
22. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
23. K. Elissa, “Title of paper if known,” unpublished.
24. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
25. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
26. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
27. Annual Conf. Magnetics Japan, p. 301, 1982].
28. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
29. Annual Conf. Magnetics Japan, p. 301, 1982].
30. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
31. Annual Conf. Magnetics Japan, p. 301, 1982].
32. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**