Report

Data Science Project

Reverse Vending Machine

Realised by

Behija Ben Ghorbel
Moatez Oueslati
Ahmed Dallali
Nour Boujmil
Cyrine Essid
Anen Halimi

Supervised by

Dorra Trabelsi

Academic year 2022-2023

Contents

1 General Introduction

The Internet of Things (IoT) and data science are playing an increasingly important role in the field of environmental sustainability.

IoT devices can collect and transmit data on environmental factors, such as air and water quality, weather patterns, and energy use. This data can then be analyzed by data scientists to better understand and monitor environmental conditions.

By using machine learning and predictive modeling, data scientists can identify patterns and trends in environmental data and make predictions about future conditions. This information can be used to develop and implement strategies to reduce waste, conserve resources, and mitigate environmental impact.

IoT and data science can also be used to monitor and optimize the efficiency of renewable energy systems, such as wind and solar power. Overall, the combination of IoT and data science is a powerful tool for promoting sustainability and protecting the environment.

2 IBM Master Plan methodology

IBM Master Plan methodology is a comprehensive and flexible approach that covers the entire project lifecycle, from initiation to closure. The methodology is based on the principles of project management and is designed to help organizations plan, execute, and deliver successful technology projects.

The IBM Master Plan methodology consists of several phases, including project initiation, project planning, project execution, and project closure. During each phase, the methodology provides guidance on key activities, such as stakeholder engagement, risk management, resource allocation, and project monitoring and control. The methodology also includes templates and tools for documenting project plans and tracking progress. The IBM Master Plan methodology is designed to be flexible and customizable, allowing organizations to adapt it to meet their specific needs and requirements.

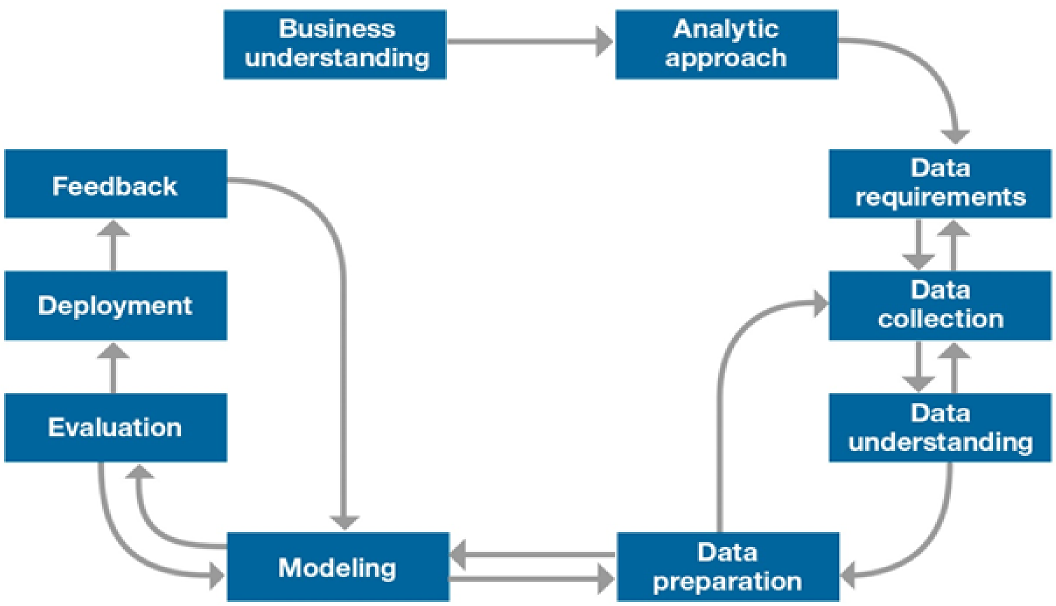


Figure 2.1: IBM Master Plan methodology

3 Business Understanding

3.1 Project context

3.1.1 Business Domain

A reverse vending machine is a device that accepts used consumer products such as bottles, cans, and plastic containers, and in exchange, dispenses a monetary or non-monetary reward, such as a refund or a coupon. The machine then sorts and compacts the collected materials for recycling. They are commonly used in supermarkets and other retail locations to promote recycling and reduce waste.
Reverse vending machines were invented to encourage people to recycle and reduce waste. By offering a reward for returning used products, the machines provide an incentive for individuals to properly dispose of their waste, rather than throwing it away. Additionally, the machines sort and compact the collected materials, making it easier for recycling facilities to process and recycle them. By increasing the amount of material that is recycled, reverse vending machines can also help to conserve natural resources and reduce the environmental impact of waste disposal.

3.1.2 Problematic

The lack of effective recycling methods has led to widespread pollution, as increasing amounts of waste continue to accumulate in our communities and natural habitats, threatening the health of both the environment and human populations. In fact,to continue with this concept, we can cite the following problems:
•**Litter and pollution:** A significant amount of litter like plastic bottles, cans, and other consumer products were not properly disposed of and ended up as litter in streets and oceans, causing pollution and environmental damage.

•**Limited recycling:** Traditional recycling methods were not always effective at collecting and processing consumer products for recycling. Many products were sent to landfills, and many of the materials were lost as a result.

•**Natural resources depletion:** The continued use of natural resources to produce new consumer products, rather than recycling existing materials, led to a depletion of resources and increased environmental impact.

•**Government regulations and pressure:** Governments around the world began to recognize the need for more effective recycling methods, and began to enact regulations and policies to encourage recycling and reduce waste.

•**Economic incentives:** Reverse vending machines provided an economic incentive for individuals to recycle, by offering a reward in exchange for returning used products. This helped to increase participation in recycling programs.

•**Technological advancements:** Advancements in technology, such as barcode and RFID scanning, made it possible to develop machines that could effectively sort and process recycled materials, making it easier to recycle large volumes of consumer products.

Overall, reverse vending machines were developed as a solution to the problem of waste and pollution caused by the lack of effective recycling methods, and to help conserve natural resources and reduce environmental impact.

3.2 Functional requirements

3.2.1 Business requirements

The business requirements for a reverse vending machine project may vary depending on the specific goals of the business. However, some common business requirements for a reverse vending machine include:

- **Item Recognition:** Reverse vending machines use computer vision and machine learning algorithms to identify and sort different types of recyclable materials, such as plastic bottles, aluminium cans, and glass. These algorithms need to be trained on large datasets of images and need to be able to recognize a wide range of materials with a high degree of accuracy.

- **Material sorting and compaction :** Material sorting involves the identification and separation of different types of materials, such as glass, plastic, or aluminum, which may be present in the beverage container. This is usually done using sensors or other detection methods that can identify the material based on its properties. Once the material has been identified, it is sorted into the appropriate compartment or container within the machine.
  Compaction is the process of compressing the material to reduce its volume and increase its density. This is done to make it easier to transport and recycle the material. The reverse vending machine may use a compaction mechanism, such as a hydraulic press, to compact the material inside the machine

before it is collected and transported to a recycling facility..

- **Deposit calculation :** Deposit calculation is the process of calculating the refund value that a customer will receive when they deposit their empty beverage container into the machine.

- **Face recognition:** In a reverse vending machine , the customer would approach the machine and stand in front of a camera. The camera would capture an image of the customer's face, which would then be compared to a database of known faces or biometric templates stored within the machine. The database might include images of registered users or previous customers who have used the machine.

- **User behavior analysis:** Reverse vending machines collect data on user behavior, such as the types of materials deposited and the frequency of use. Data science techniques such as data visualization and statistical analysis can be used to understand user behavior and improve the design of the machines.

- **Customer rewards and vouchers:** Reverse vending machines provide rewards or vouchers to customers who deposit recyclable materials. Data science techniques such as machine learning can be used to optimize the rewards program and personalize the rewards based on individual user behavior.

These are some of the key business requirements for a reverse vending machine project, but specific requirements may vary depending on the business goals, target audience, and other factors.

3.2.2 Data Science requirements

- **Data augmentation:** Data augmentation is a technique used to artificially increase the size of a dataset by creating new training examples from existing ones. The idea behind data augmentation is to make slight modifications to the original data so that the model can learn to generalize better and become more robust to variations in the input data.

- **CNN:**Item recognition with Convolutional Neural Networks (CNN) is a technique used to classify and identify different items in an image. CNNs are particularly well-suited for this task because they can effectively extract features from images and learn to identify patterns in the data.

- **OCR**:Text extraction with Optical Character Recognition (OCR) is a process that involves automatically detecting and extracting text from images or scanned documents. OCR technology uses computer algorithms to analyze the pixels in an image and identify characters, words, and other textual elements.

- **OpenCv**: Face recognition with OpenCV is a technique that uses computer vision algorithms to automatically detect, recognize, and verify faces in images or video streams. OpenCV is a popular open-source computer vision library that provides a range of tools and algorithms for image processing, feature detection, and machine learning.

.

3.3 Non-functional requirements

Non-functional requirements for a reverse vending machine project might include:

• Security and data protection measures to ensure that personal information of the user is protected.

- Durability and reliability to withstand heavy usage and harsh environments.
- Energy efficiency to minimize electricity consumption.
- Accessibility for users with disabilities, such as visual and auditory interface.
- Scalability to adapt to changing demand and product types.

3.4 Technical constraints

The development of a reverse vending machine requires advanced technical capabilities, including machine learning, computer vision, and robotics. These technologies may be difficult to implement and may require specialized expertise.

- Cost constraints: The development and implementation of a reverse vending machine can be expensive, especially if the machine is required to have advanced features such as real-time monitoring and item recognition.

- Environmental constraints: Reverse vending machines are usually placed in public areas and need to operate in various weather conditions, which can put constraints on their design, such as waterproofing, cooling and heating systems.

- Regulatory constraints: There may be regulations and laws that need to be followed when developing and implementing a reverse vending machine. For example, safety standards and environmental regulations.

- Maintenance and repair constraints: Reverse vending machines require regular maintenance and repairs to ensure that they continue to operate correctly. This may include regular cleaning and replacement of parts.

- Data privacy and security constraints: Reverse vending machines collect data from users, which must be kept secure and private. This requires implementing security measures such as encryption and access controls to protect sensitive information.

- Integration constraints: Reverse vending machines need to integrate with existing systems, such as inventory management, customer rewards, and financial systems, which can be difficult to implement and may require significant technical resources.

4 Data Collection

Data collection for a reverse vending machine (RVM) involves gathering information about the machine's usage, performance, and customer behavior. The data can be used to improve the RVM's efficiency, identify opportunities for cost savings, and enhance the customer experience.

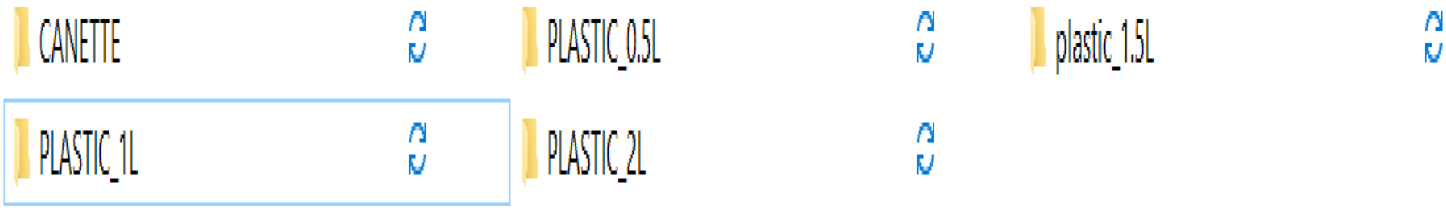


Figure 4.1: DATA COLLECTION

5 Data Understanding

To develop a Support Vector Machine (SVM) model using the RVM algorithm, we need to have a good understanding of the data we are working with. Here are some steps we can take to understand our data:

5.1 Data exploration:

We should begin by exploring our data to get a better sense of its characteristics. This can include visualizing our data, identifying any patterns or trends, and checking for outliers or missing values.

5.2 Feature engineering:

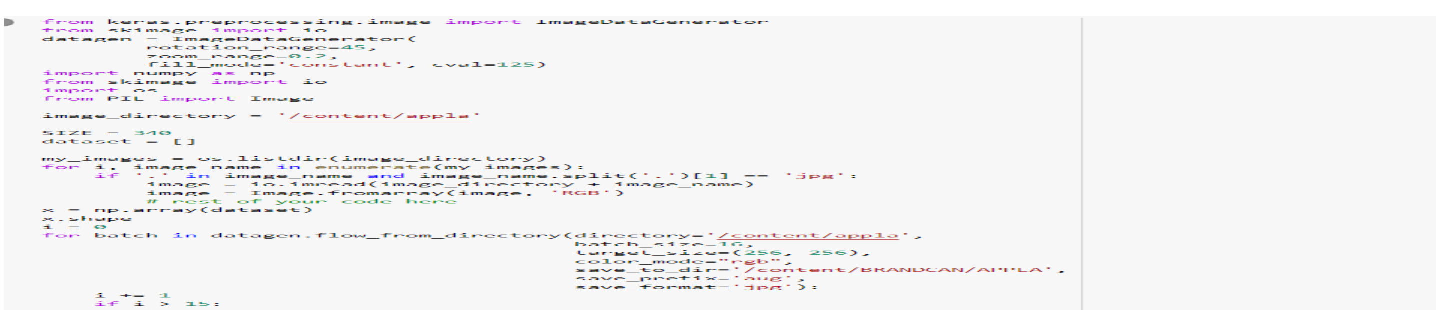
Feature engineering involves transforming our data into a set of features that can be used by the RVM algorithm. This can include creating new features, transforming existing features, or selecting a subset of the available features.

By following these steps, we can develop a good understanding of our data and build an effective RVM model.

6 Data preparation

6.1 Data augmentation:

Data augmentation involves generating additional training data by applying random transformations to existing images, such as flipping, rotating, zooming, or cropping. This can help the model generalize better to new images and reduce overfitting.



```
from keras.preprocessing.image import ImageDataGenerator
from skimage import io
datagen = ImageDataGenerator(
        rotation_range=45,
        zoom_range=0.2,
        fill_mode='constant', cval=125)
import numpy as np
from skimage import io
import os
from PIL import Image

image_directory = '/content/appla'

SIZE = 340
dataset = []

my_images = os.listdir(image_directory)
for i, image_name in enumerate(my_images):
    if '.' in image_name and image_name.split('.')[1] == 'jpg':
        image = io.imread(image_directory + image_name)
        image = Image.fromarray(image, 'RGB')
        # rest of your code here
x = np.array(dataset)
x.shape
i = 0
for batch in datagen.flow_from_directory(directory='/content/appla',
                                          batch_size=16,
                                          target_size=(256, 256),
                                          color_mode="rgb",
                                          save_to_dir='/content/BRANDCAN/APPLA',
                                          save_prefix='aug',
                                          save_format='jpg'):

    i += 1
    if i > 15:
        break
```

Figure 6.1: DATA AUGMENTATION

6.2 Resizing and normalization:

Images may need to be resized to a standard size so that they can be easily processed by the model.

```
[ ] from keras.preprocessing.image import ImageDataGenerator

    #We need to normalize image
    datagentrain = ImageDataGenerator(featurewise_center=True, featurewise_std_normalization=True,rescale=1./255)
    datagentest = ImageDataGenerator(featurewise_center=True, featurewise_std_normalization=True,rescale=1./255)
```

```
[ ] datagentrain = ImageDataGenerator(featurewise_center=True, featurewise_std_normalization=True,
                                      brightness_range=[0.1,1.0]
                                     )
    datagentest = ImageDataGenerator(featurewise_center=True, featurewise_std_normalization=True)
    datagentrain.mean = np.array([127.5, 127.5, 127.5], dtype=np.float32).reshape((1,1,3))
    datagentrain.std = 255.0
    datagentest.mean = np.array([127.5, 127.5, 127.5], dtype=np.float32).reshape((1,1,3))
    datagentest.std = 255.0
```

Figure 6.2: NORMALIZATION

6.3 Data splitting:

Before we begin training our RVM model, we should split our data into training and testing sets. The training set will be used to train our model, while the testing set will be used to evaluate its performance.

```
[ ] from sklearn.model_selection import train_test_split
    train_df, test_df = train_test_split(df,test_size = 0.25, stratify=df.label)
```

```
[ ] traingenerator = datagentrain.flow_from_dataframe( train_df , './',x_col = 'image' , y_col = 'label',    target_size=(340, 768),
                                                       batch_size=32,
                                                       class_mode='categorical')
    testgenerator = datagentest.flow_from_dataframe( test_df , './',x_col = 'image' , y_col = 'label',    target_size=(340, 768),
                                                       batch_size=32,
                                                       class_mode='categorical')
```

Figure 6.3: DATA SPLITTING

7 Modeling and Evaluation

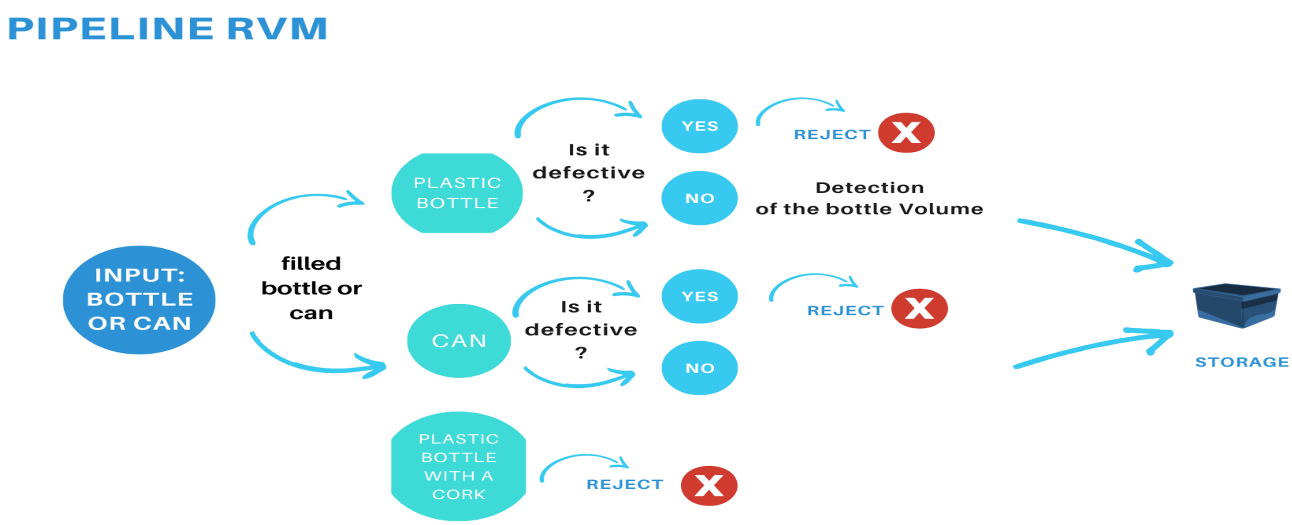


Figure 7.1: PIPELINE

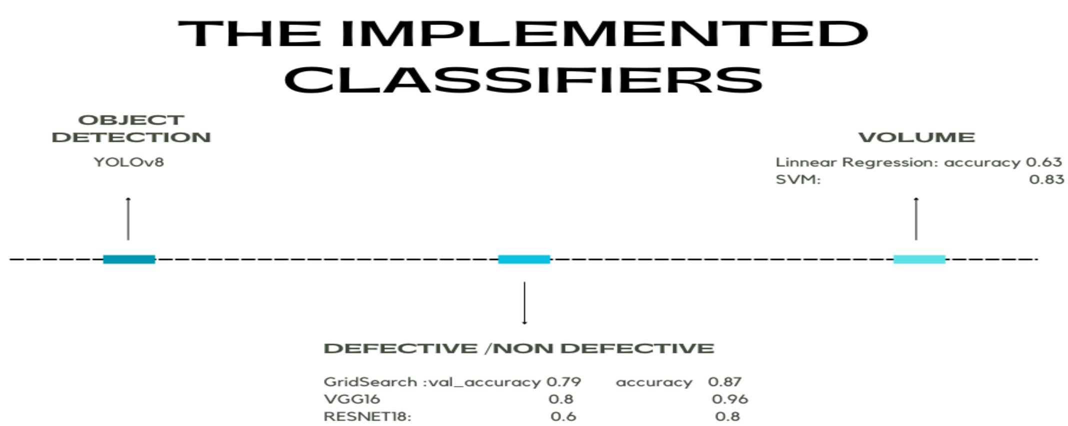


Figure 7.2: The implemented classifiers

7.1 Object Detection :

We have utilized YOLOv8 for object detection in order to identify whether a bottle is a can or a plastic bottle, and whether it has a cap or not.

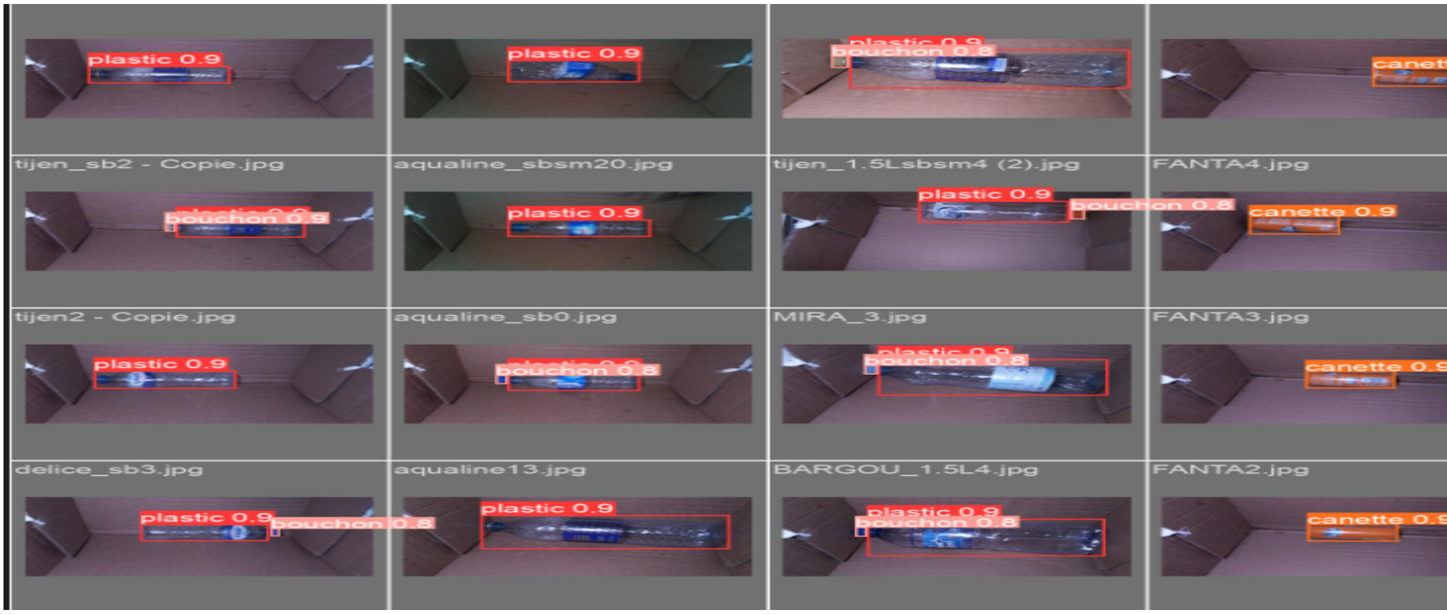


Figure 7.3: Object Detection

7.2 Defective / Non-Defective :

A deep learning model can be used to identify whether a bottle is defective or not using convolutional neural networks (CNNs) with pre-trained models such as VGG16 and ResNet18. Additionally, we used hyperparameter tuning techniques such as grid search to optimize the model's performance.

```
[ ] predict("/content/Capture.PNG")
```


```
1/1 [==============================] - 0s 23ms/step


Maximum Probability:  0.67244744
Classified: defective



------------------Individual Probability-------------------------------

DEFECTIVE : 67.24 %
NON-DEFECTIVE : 32.76 %
```

Figure 7.4: Defective


```
1/1 [==============================] - 6s 6s/step


Maximum Probability:  0.81996953
Classified: non-defective



-------------------Individual Probability--------------------------------

DEFECTIVE : 18.0 %
NON-DEFECTIVE : 82.0 %
```

Figure 7.5: non-Defective

7.3 Volume :

To determine the volume of the bottle, we used OpenCV to extract bottle dimensions into a CSV file. We then applied a Support Vector Machine (SVM) model as the first method and used YOLOv3 to extract dimensions into another CSV file as the second method. We applied two models, SVM and Linear Regression, to the extracted data and found that SVM had better accuracy compared to Linear Regression.

```
# Utiliser le modèle pour prédire la classe d'une nouvelle bouteille
new_bottle = "/content/drive/MyDrive/RVM2/TEST/PLASTIC_1L/mira_sb2.jpg"
dim_new_bottle = calculate_bottle_volume(new_bottle,net)
print(dim_new_bottle)
predicted_class = svm.predict([dim_new_bottle])
print(f"Classe prédite : {predicted_class}")
```

```
[809, 188]
Classe prédite : ['PLASTIC_1L']
```

Figure 7.6: SVM

```python
# Utiliser le modèle pour prédire la classe d'une nouvelle bouteille
new_bottle = "/content/drive/MyDrive/RVM2/TEST/PLASTIC_1L/mira_sb2.jpg"
dim_new_bottle = calculate_bottle_volume(new_bottle,net)
print(dim_new_bottle)
predicted_class = model.predict([dim_new_bottle])
print(f"Classe prédite : {predicted_class}")
```

```
Exactitude du modèle : 0.6346737951104092
[809, 188]
Classe prédite : [0.6797179]
```

Figure 7.7: Linear Regression

8 Conclusion

The implementation of reverse vending machines, combined with data science, has the potential to revolutionize waste management by providing a convenient and efficient solution for individuals to recycle, while also improving waste collection and processing through the use of data analysis.

In fact, by using data to track usage patterns, assess the impact of recycling programs, and identify areas for improvement, organizations can make informed decisions that lead to more sustainable waste management practices.

Additionally, the use of reverse vending machines can increase public engagement and participation in recycling programs, further promoting a culture of sustainability and environmental responsibility

Overall, the combination of reverse vending machines and data science can play a key role in promoting a more circular economy and a cleaner, healthier planet.