Group - 10

VIDA

MAKING AN ACTIVE EFFORT

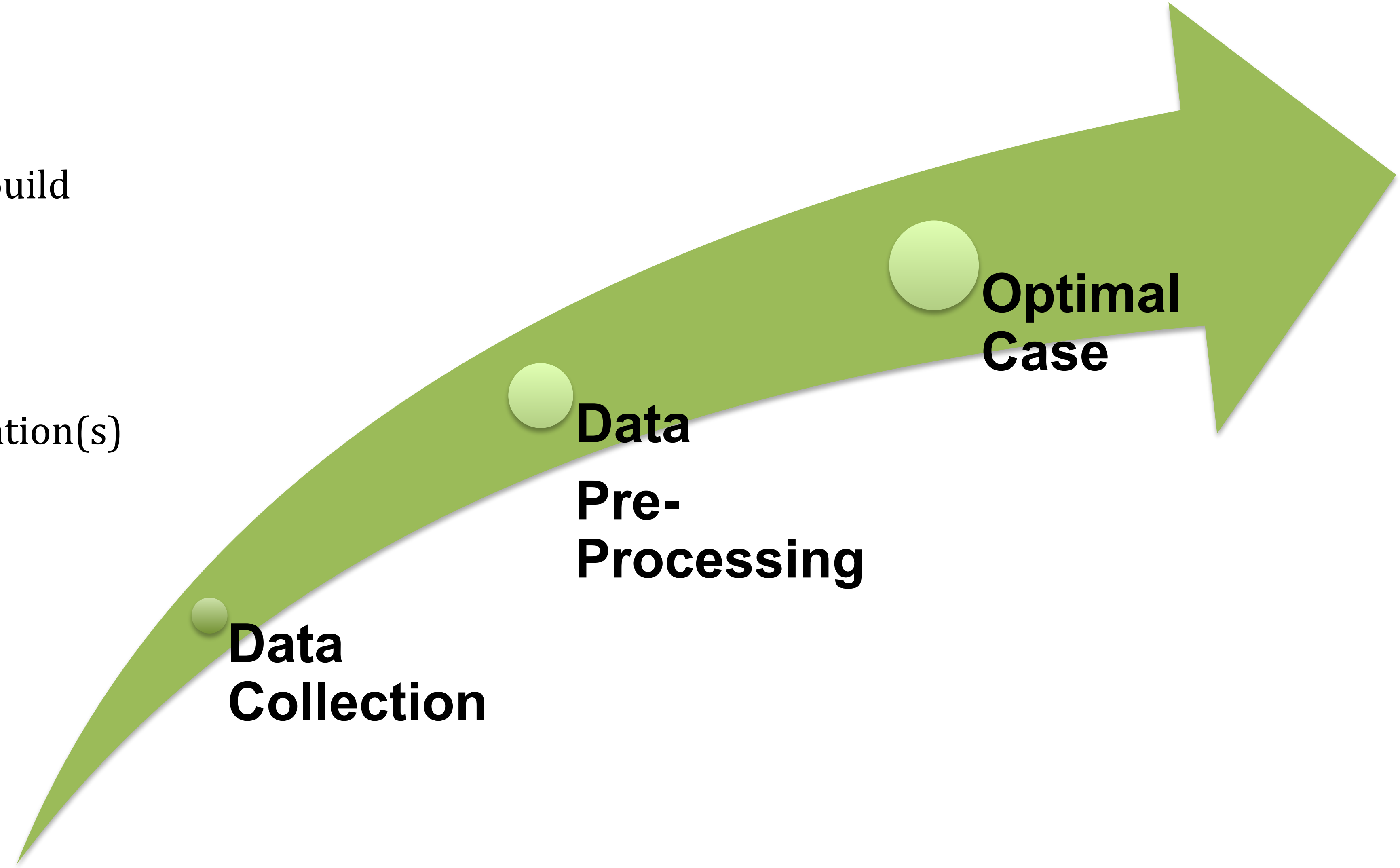
Project Outline

- **Waste management** is a major issue because of the volume of waste produced daily in the world. Waste classification is a big aspect of this issue. But what if we could automate the process using AI?
- To manage a spectrum of waste products, it's crucial to have an **advanced/intelligent** waste management system.
- The process of separating the waste into its many components is one of the most crucial parts of waste management, and it is typically carried out manually by hand-picking.
- To simplify the process, we propose an intelligent waste material **classification system.**

CORPORATE
GURUKUL

Index

I. Team members
II. What are we trying to build
III. Abstract
IV. Concepts Used
V. Algorithm Used
VI. Comparison of models
VII. Diagram(s) or Visualization(s)
VIII. Final Recommendation
IX. Business use cases
X. Challenges Faced
XI. Our Learning
XII. Future Enhancements
XIII. Demo



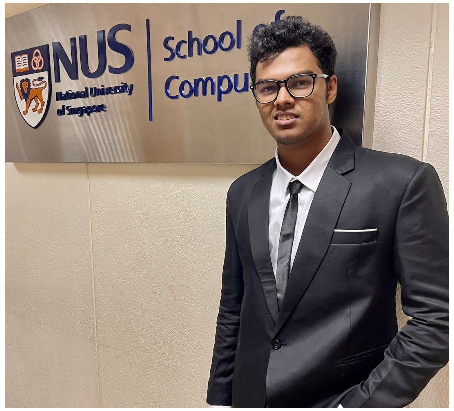
CORPORATE GURUKUL

We the Team!

TEAM LEAD : JOHN HARSHITH
MODEL BUILDING & DEPLOYMENT

ANITA PRIYADARSHINI A
FRONTEND DEVELOPMENT & DEPLOYMENT

J SAKETH KUMAR
DATA PROCESSING &
HYPERPARAMETER TUNING



ISAAC RITHARSON P
MODEL BUILDING & EVALUATION (USING ORANGE)

ADITI DUDHORIA
BUSINESS PLANNING &
PRESENTATION

PRIYA TIWARI
DATA ANALYSIS &
HYPERPARAMETER TUNING

CORPORATE
GURUKUL

What we are trying to build?

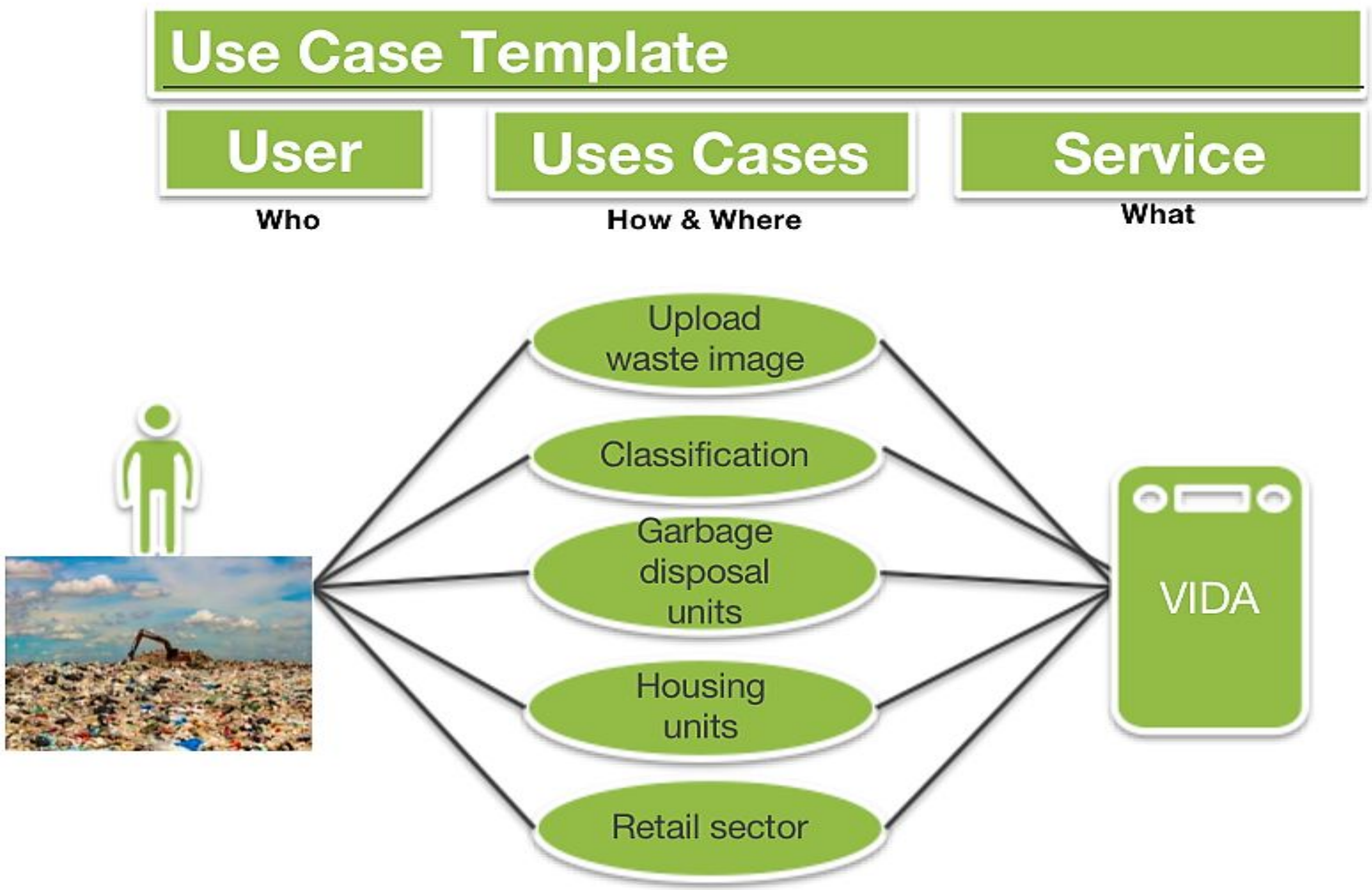
- We are attempting to build a web application that can help users easily segregate waste into one of **6 classes**:- Cardboard, Glass, Metal, Paper, plastic, and trash.
- We have used deep learning models to classify our uploaded image.
- Besides this, we have implemented a **web interface** to accompany the same.

Abstract

- The classification of trash is essential for **resource recycling** and environmental protection. Different categories of garbage are now defined by laws passed in various nations on garbage classification.
- However, it was discovered during the implementation of these laws that were accurately differentiating between various sorts of waste still needs to be solved.
- In this project, the work of garbage classification will be finished using a **deep learning model**. In particular, the classification performance of **VGG16** and **CNN** based on a publicly accessible image data set is compared.
- We discovered that the overall method's prediction results are more precise than those of a single neural network model and among the many ensemble strategies.
- The model with the **highest accuracy** is chosen and deployed with the Web interface.

CORPORATE
GURUKUL

Business Use Case



CORPORATE GURUKUL

BUSINESS MODEL CANVAS: VIDA

KEY PARTNERS

Plastic produce manufactuers

National and Regional Governments and their corportaions like municipalities.

FMCG and pharma companies

Medium and Big Scale industries

Retail sector

Academic establishments

Food delivery companies like

Zomato, Grab

KEY ACTIVITIES

Segregation of wastes for reuse and recycle.
take- back management
Waste collection and it's analysis
Sorting process optimization

KEY RESOURCES

Collection of cardboard, glass, metal, paper, plastic, and trash.
Sensors and imaging
Mainetence and management
Conveyor Belts
Installation team

VALUE PROPOSITIONS

Environmental:Rational model management
Economic: Energy savings and tapping on the big plastic market.
Social: Improved quality of life (Less plastic pollution and easier management of reuseable waste)

CUSTOMER RELATIONSHIP

Community based
Long term

CHANNELS

Through VIDA application
Refferal
Outreach
Social Media and through mails

CUSTOMER SEGMENTS

Big scale industries who needs to segregate their waste.

Entreprenuers who wants to promote sustainiblity through their brand.

Institutions like schools, colleges who wants to teach reuse of waste

FMCG, pharms, retails companies who wants to use this for further packaging.

COST STRUCTURE

Application and installation charges
Upkeep and maintainence
Salaries and training people
Marketing and promotions
Taxes

Revenue Stream

Revenue from the subsciption model (individual, businesses and educational
Installation and mainetence (pay per service model)

Important Use Case

- Most standard business models that revolve around the utilizing **waste as a resource** are:
  - companies that collect and transport the waste
  - companies that treat waste after separating it
  - companies that produce energy from waste
  - companies that sell the waste after processing it into other products
- Our focus is on the second case model, i.e., the separation of wastes.
- The population has increased, and with it, waste generation has grown. Manual segregation leads to a **waste of time** and **money**. So our most important use case solves this problem for our target audience by using a deep learning model to do this work for you efficiently. We can further improvise this process by **automatically** capturing the photos as the waste moves through a conveyor belt or any similar mechanism. This can significantly reduce human intervention in trash segregation and improve the accuracy of the same.

Concepts Used

Image Classification

- Data Scraping (Collection)
- Data Cleaning
- Data Augmentation
- Data Preprocessing (Resizing)
- Train-Test Split for Training and Testing
- Model selection
- Model Building/Training (Tensorflow/Keras)
- Testing and Evaluation
- Model Optimization and Hyper Parameter Tuning
- Final Results
- Interpreting Visualizations (Orange Data Mining)

orange
DATA MINING

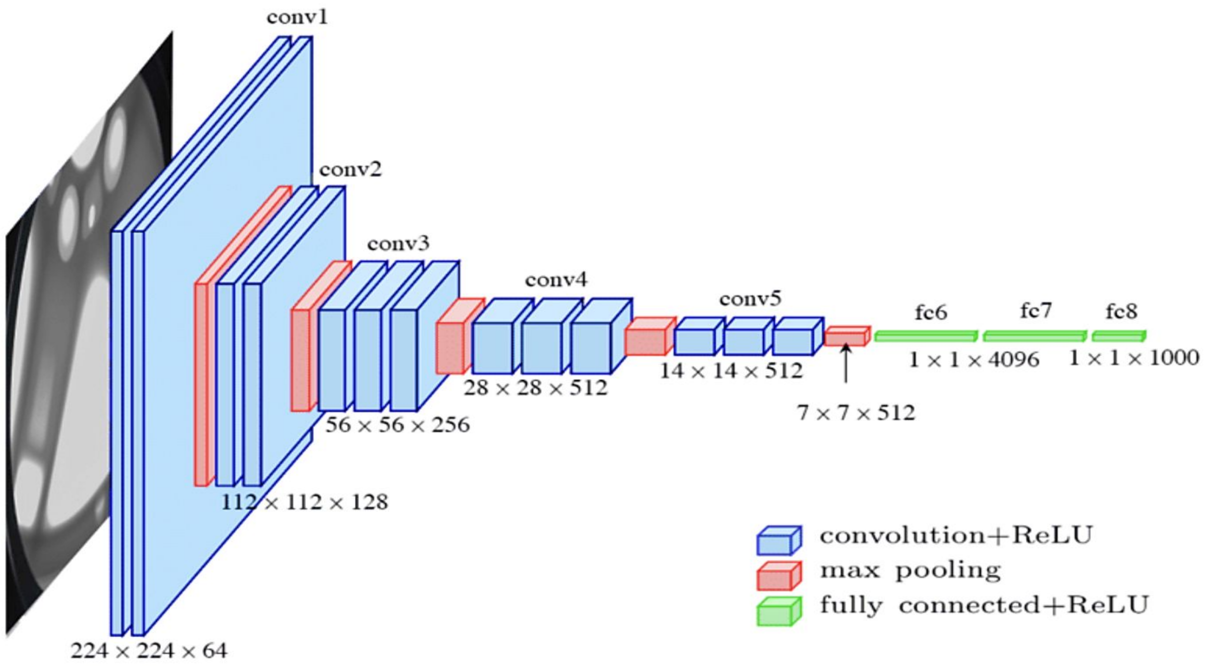
+ Keras

CORPORATE
GURUKUL

Algorithm Used

For our model the algorithms we've used are:

- **CNN** (Convolutional Neural Network)
- **VGG-16**(Visual Geometry Group-16
)



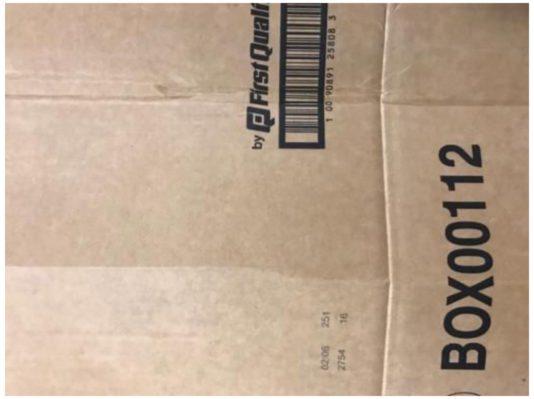
CORPORATE
GURUKUL

Our Data

About Dataset

The data was obtained through the Kaggle website and The Garbage Classification Dataset contains 6 classifications: cardboard (393), glass (491), metal (400), paper(584), plastic (472) and trash(127).

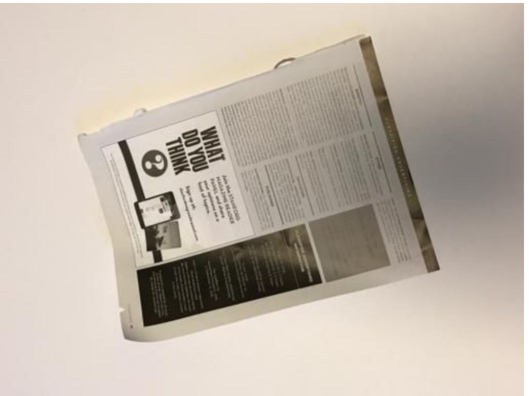
1.cardboard



2.glass

3.metal

4.paper



5.plastic

6.trash

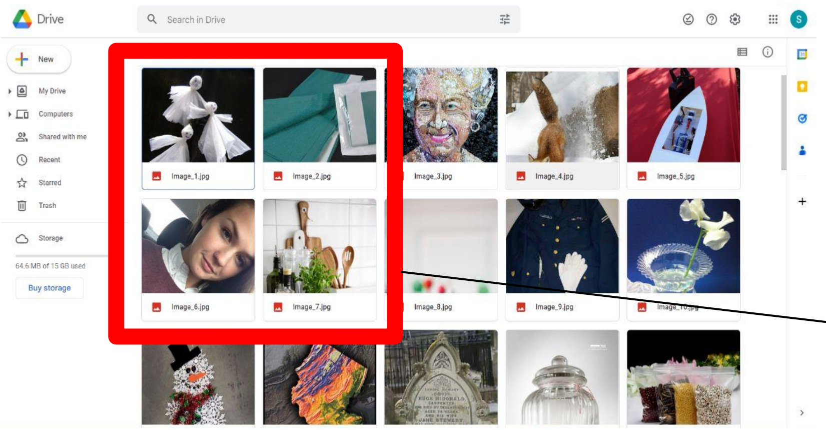
CORPORATE
GURUKUL

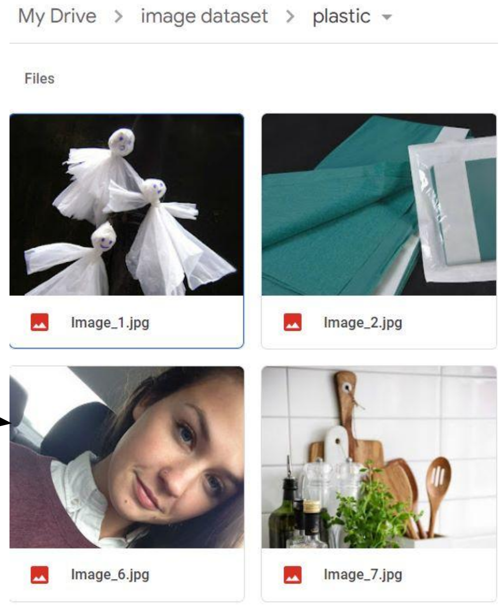
Data Collection

- Since we require an image dataset, we had to scrape images from the internet.
- For web-scraping we used an **API**.
- Besides scraping images from the internet, we even took garbage images from Kaggle.
- We were able to scrap **813 images** from the internet and able to get **1982 images** from **Kaggle**.
- The whole dataset was then divided into six categories, namely : **Cardboard, glass, plastic, metal, paper and other garbage.**

Data Pre-processing

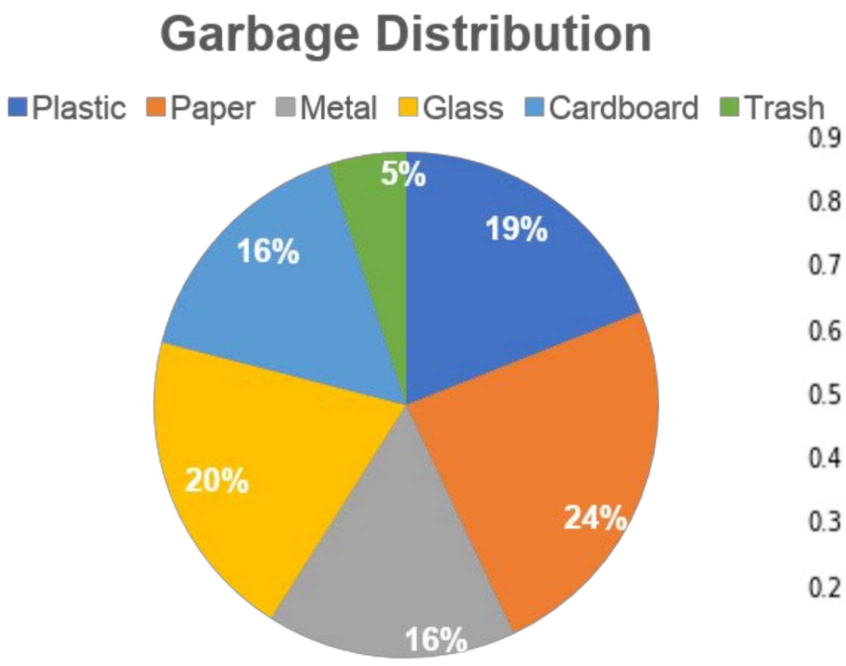
- Though we had the dataset, the images from Kaggle were relatively much more cleaner than compared to the web-scrapped ones, so the images from web had to be manually cleaned, since many contained **irrelevant images**.
- Finally after clearing 268 images from our dataset, we were left with **2527 relevant images**.
- As far as the processing of the images goes, after cleaning our dataset, we then had to **resize** all the images of the dataset, so as to get an uniform sized images in our dataset.
- Then we **augmented** our data to reduce overfitting by forming new and different examples to train datasets.

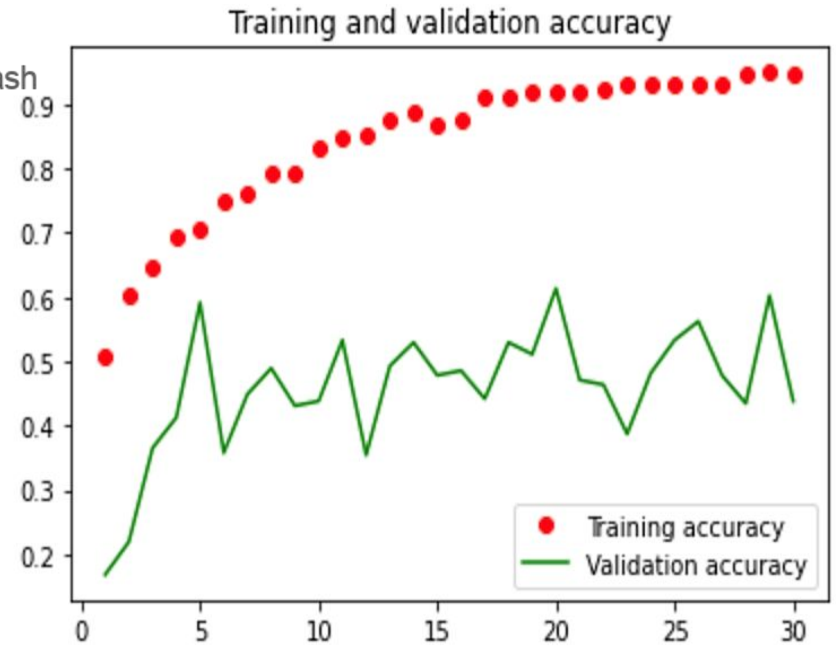


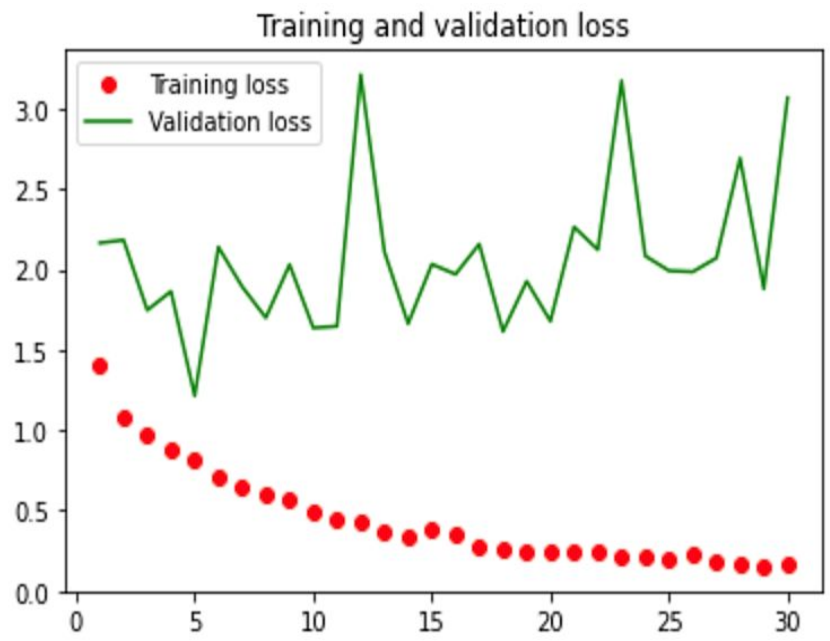


CORPORATE GURUKUL

Diagram(s) or Visualization(s)







CORPORATE GURUKUL

Optimal Case

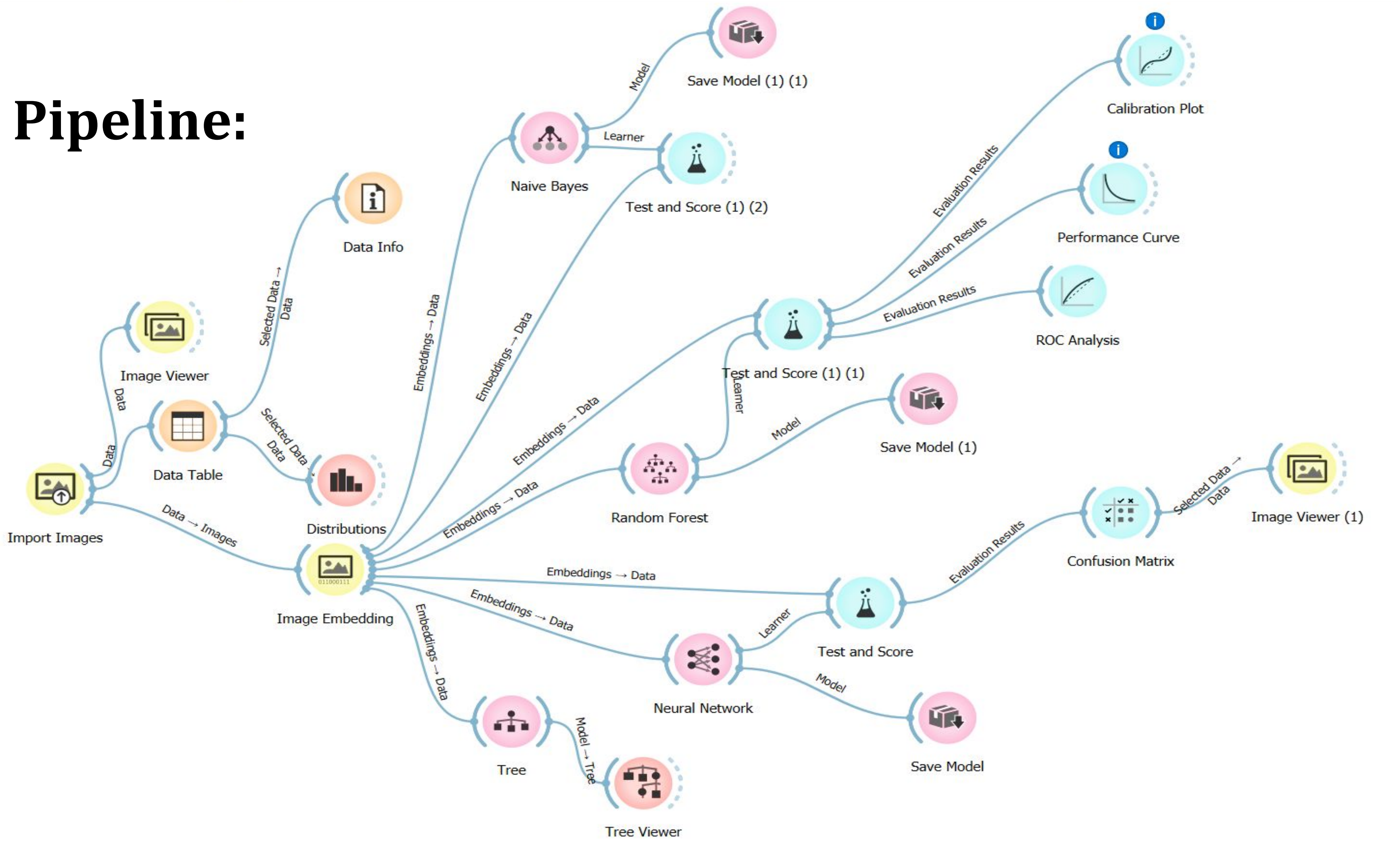
- **Training Accuracy ~ 97%**
- **Test Accuracy ~ 95%**

Hyperparameter :
Number of epochs = 100
Loss function = Categorical cross entropy
Optimizer function = Adam
Train: Test split = 0.2

Using the above parameter values, will give us the optimal model (VGG16 in this case here)

Validation using Orange Software

Model Pipeline:



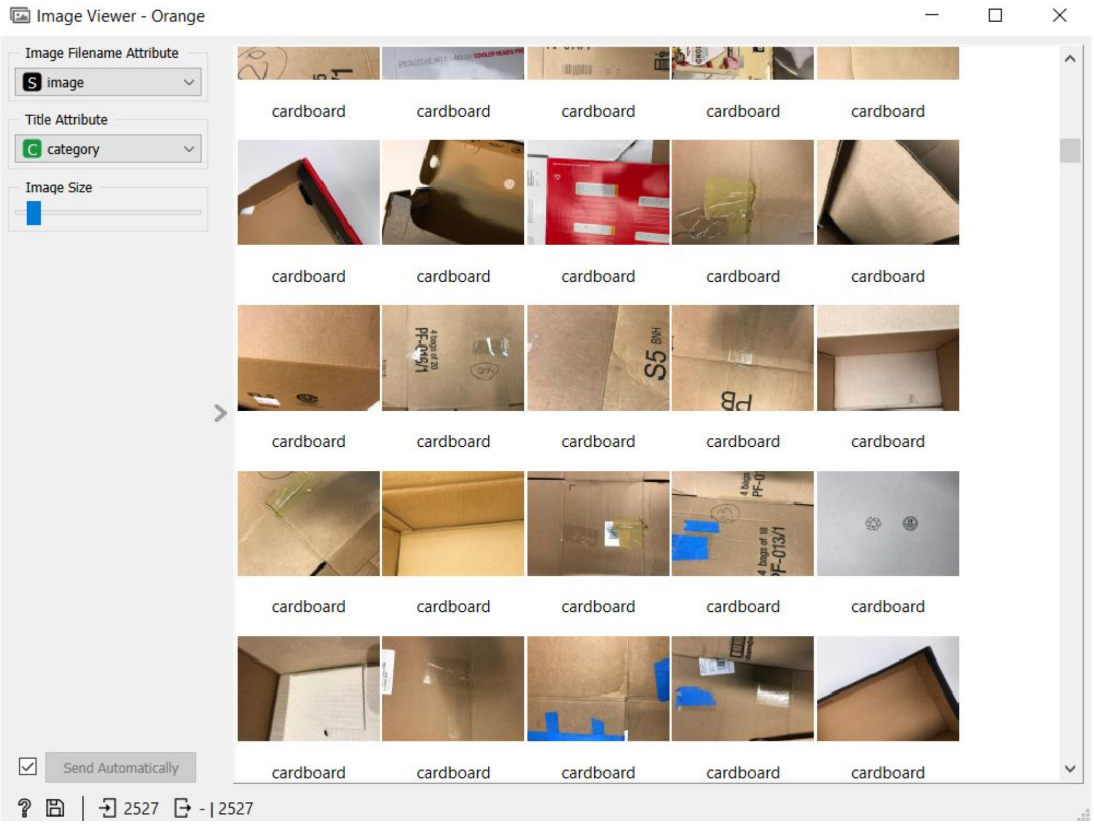
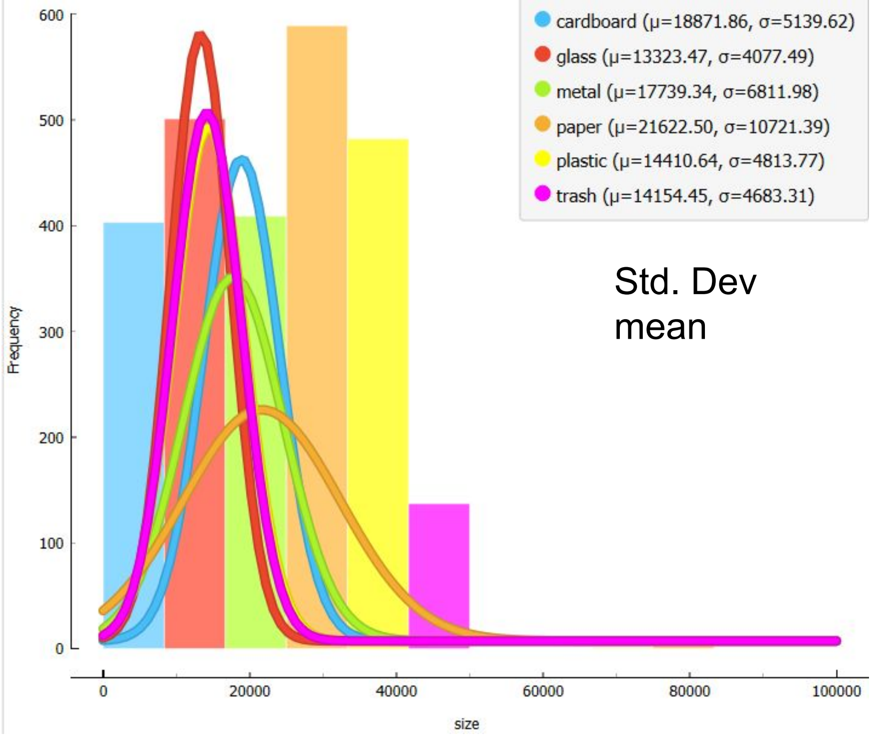


Image Viewer



Std. Dev
mean

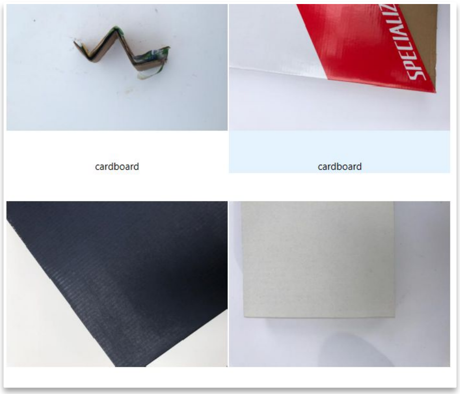
Size Distribution of Images

Actual	cardboard	glass	metal	paper	plastic	trash	Σ
cardboard	369	1	1	27	1	4	403
glass	0	435	21	3	40	2	501
metal	3	25	352	5	20	5	410
paper	10	3	5	559	2	15	594
plastic	0	34	21	8	404	15	482
trash	3	1	16	21	17	79	137
Σ	385	499	416	623	484	120	2527

MPL Model with 100 hidden neurons, Activation= ReLu, Optimizer = Adam , Alpha(Learning Rate =0.0001), Max_iterations = 200

CORPORATE GURUKUL

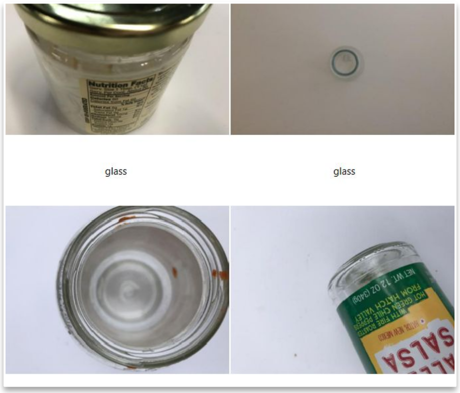
Actual- glass
Predicted-plastic (40 images)



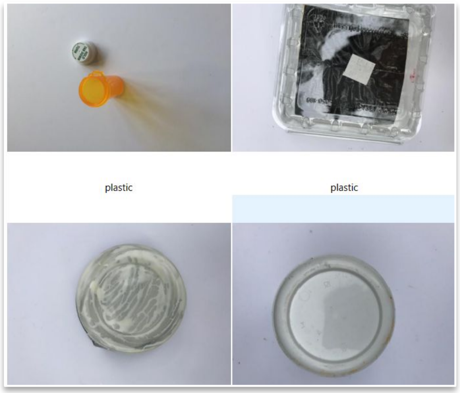
Actual- Cardboard
Predicted-Paper (27 images)

Actual- plastic
Predicted-glass (34 images)

Actual- glass
Predicted-plastic (40 images)



Actual- Glass
Predicted-Metal (21 images)



Actual –Plastic
Predicted- Metal (21 images)

CORPORATE
GURUKUL

Comparative Analysis

Model	AUC	CA	F1	Precision	Recall
MLP	0.982	0.870	0.869	0.869	0.870

Model	AUC	CA	F1	Precision	Recall
Random Forest	0.911	0.708	0.698	0.704	0.708

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.946	0.724	0.742	0.782	0.724

Model	AUC	CA	F1	Precision	Recall
Tree	0.745	0.586	0.586	0.586	0.586

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

$$ificity = \frac{TN}{TN + FP}$$

To be able to prove that Deep Learning Models Performed Better than Machine Learning on our Dataset

CORPORATE GURUKUL

Final Recommendation

- Using transfer-learning models like ResNet, MobileNet etc could have led to slightly better accuracy than the ones we used. This is because by using pre-trained model as a starting point, it is possible to leverage the knowledge learned by the model and improve the performance of the new model on the task at hand. **VGG-16** gave the highest accuracy otherwise.
- The model needs to be **trained more** on a **larger dataset** to get better results with an accuracy that can be deployed in the industrial level.
- The number of classes needs to be increased to include various types as **subclass** in the classes itself. For eg: plastic could be further be polythene, pvc etc.

Challenges Faced

- During the initial stages of the project, all members of the team had to **learn** and understand the techniques used in implementation of deep learning in real world applications and the approaches taken to successfully build a business concept/model from the problem statement
- We had to thoroughly **understand** the data science life cycle and how it worked from data acquisition to deployment of models in order bring an actionable useful outcome
- During the Data preprocessing step, we faced difficulty in exploring the dataset containing several images of the garbage waste. Later while building the model we had to choose the models which were quite efficient in performance in terms or accuracy, training and time period.
- Finally building an interactive user website and **deploying** it on Azure was quite challenging since we were all new to the cloud interface.

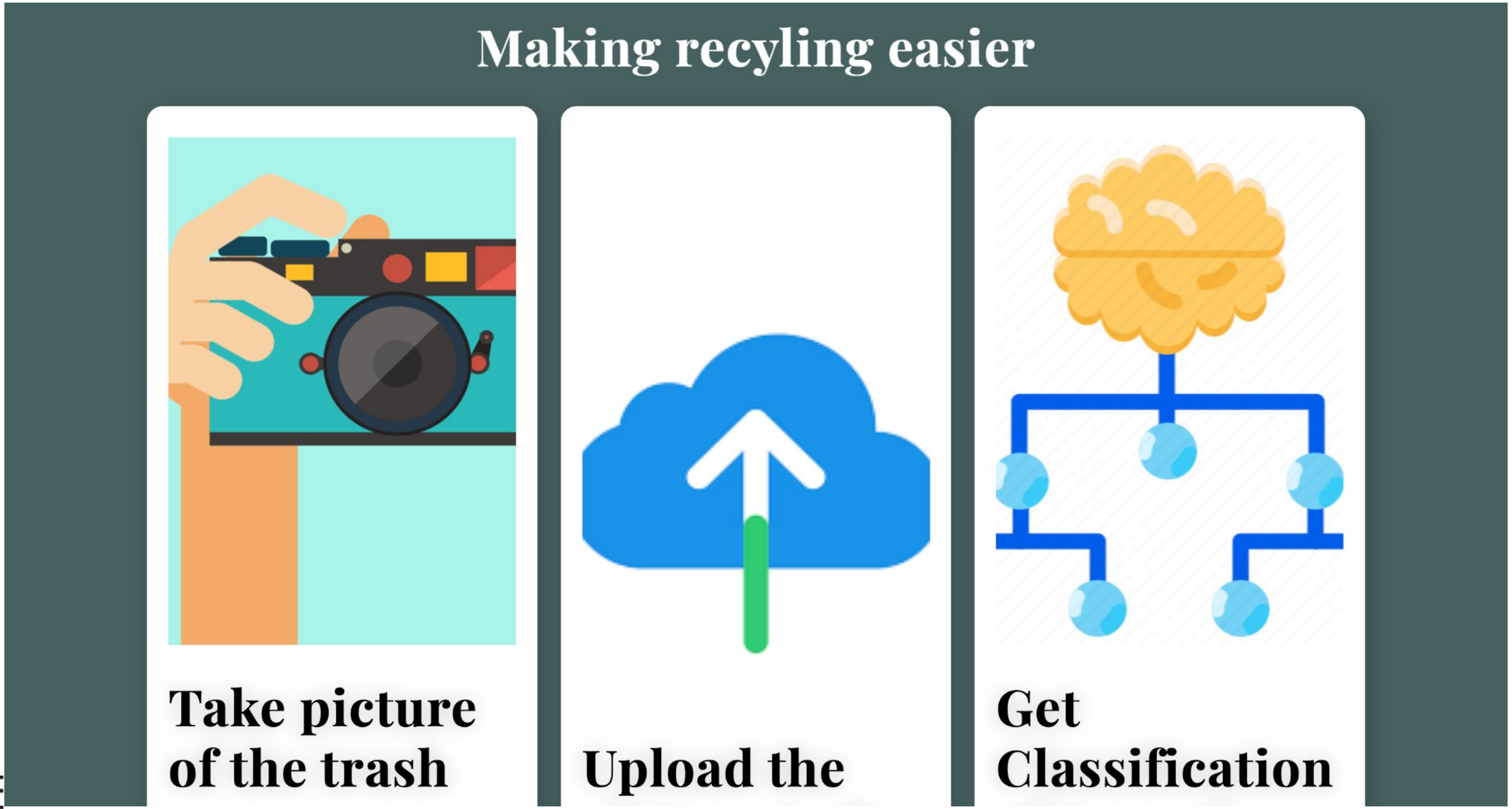
Our Learning

- Through this project, we understood the importance of an **outcome-based approach** in a data science life cycle and the impact of Deep Learning techniques to improve the efficiency and performance of humankind in day-to-day activities.
- We also learned the in-depth concepts of convolutional neural networks and their implementation in python, along with several models like VGG-16, etc. We implemented concepts of **Flask** and **Azure**.
- It was interesting to see a solution to a real-world problem and how it could potentially create a massive market in the industry and contribute to the social community and its welfare.
- There were few instances where we had to scrap our models and start the work again due to constraints about models which were not allowed to use. We tried our best to fight these circumstances **together as a team**.
- It was also great fun to work as a team and understand several perspectives of the different members on the same concept, which helped everyone share their knowledge and **help each other** out in difficult situations (i.e., becoming team players).

Future Enhancements

- More models will be put to the test for the same purpose, and we'll determine which model performs best before implementing it.
- To further ensure substantially larger accuracy, we are intending to drastically expand the size of the dataset.
- Improve the interface of the webpage and our model, so that this model of ours is easy-to-use by the end-user.
- Actually deploying the model, so that it will be accessible to anyone with internet.
- Using transfer learning models to get a slightly better accuracy and performance.

Demo



CORPORATE GURUKUL

TEAM NO: 10

CORPORATE
GURUKUL

Thank You