MSc Final Project: Waste Classification using Deep Learning

# Objective

This project aims to develop a deep learning-based classification system to distinguish between organic and recyclable waste. The system utilizes three pre-trained convolutional neural networks—VGG16, ResNet50, and EfficientNetB0—leveraging transfer learning for image classification. The models are evaluated on performance metrics such as accuracy, AUC, precision, recall, and explainability using SHAP and LIME.

# Dataset Description

The dataset is structured into two main categories: organic (O) and recyclable (R). The directory is organized as follows:  
  
DATASET/  
├── TRAIN/  
│ ├── O/ # Organic training images  
│ └── R/ # Recyclable training images  
└── TEST/  
 ├── O/ # Organic test images  
 └── R/ # Recyclable test images

# Methodology

All three models follow a unified approach:  
- Preprocessing with data augmentation and normalization  
- Transfer learning using pretrained ImageNet weights  
- Adding custom dense layers for binary classification  
- Early stopping to prevent overfitting  
- Evaluation using accuracy, AUC, precision, recall, and F1-score  
- Visual explanation using SHAP and LIME

# Model 1: VGG16

VGG16 yielded the best results with a high AUC of 0.95 and balanced performance across precision, recall, and F1-score. Its training and validation loss curves showed consistent improvement and convergence, indicating effective learning and generalization. Explainability through LIME and SHAP confirmed the model's attention to relevant image features.

# Model 2: ResNet50

ResNet50 showed moderate performance with an AUC of 0.64. It exhibited strong recall for recyclable waste but failed significantly on organic classification. The confusion matrix revealed a bias towards recyclable labels, making it unsuitable for real-world usage without further tuning.

# Model 3: EfficientNetB0

EfficientNetB0 performed the worst, with an AUC of only 0.53—just above random chance. While the loss curve decreased slightly, the AUC plateaued with no significant learning. The model likely requires deeper fine-tuning or a larger dataset to be effective.

# Conclusion and Best Model

Among the models tested, VGG16 emerged as the most robust and reliable solution for waste classification. It combined high accuracy, effective learning curves, and strong visual interpretability, making it the best choice for deployment.