**Project 3 Proposal**

**Title: Machine Learning Waste Classification Model**

**Project Members:** Lareena Llamado, Kathy Manthey, Manuela Muñoz, Nicole Muscanell

**Background & Purpose**

The purpose of this project is to build a machine learning model that classifies images of materials and objects as being organic or recyclable. Recycling in the United States is a huge challenge. The U.S. produces more waste than any other country in the world—close to 300 million tons a year (United States Environmental Protection Agency, March 2020). Of this, only about 30 percent of the waste material is recycled.

One reason the recycling rate has been relatively low compared to other parts of the world is because the U.S. does not properly sort their waste material, instead only collecting recyclables into one bin, causing the material to be contaminated (Pahl, August 2020). One major result of this improper sorting is that in 2018, China, who historically recycled a large amount of waste from the U.S., began rejecting the waste largely because of the contamination. A second major challenge is that because there is a tremendous amount of waste to sort through, it is hard to accomplish this in a safe and efficient manner using solely human workers. Not only is it time consuming, but waste sorting has a number of risks, for example workers are often exposed to hazardous materials, and sometimes don’t have the proper protective gear necessary for handling such materials (Chen, June 2015).

One promising area of research and innovation that has begun to address this problem is the machine learning and artificial intelligence. Using smart technology to classify waste could be a cost efficient, safe, and possibly even more accurate method of sorting large amounts of waste in a timely manner, which could thereby help to improve the recycle rate.

**Technical aspects**

We will use Keras and Tensorflow to build a deep learning convolutional neural network. Our group plans to train several different models so that we can adjust several features such as the number of layers, image size, filter size, dropout, learning rates, etc. After the trained model is saved, we plan to deploy the test model as a simple web app via Python Flask and Heroku. The web app would allow users to upload their own images and receive a classification of their image(s) in real time. We will also use Python libraries (Matplotlib, Seaborn) to plot the results of the model training and validation.

**Data**

The dataset has over 22,5000 images and is divided into train data (85 %) and test data (15 %).

<https://www.kaggle.com/techsash/waste-classification-data>

**References**

Pahl, C. (2020, August 31). How Machine Learning and Robotics are Solving the Plastic Sorting Crisis. *PLUGANDPLAY. Retrieved from* <https://www.plugandplaytechcenter.com/resources/how-ai-and-robotics-are-solving-plastic-sorting-crisis/>

United States Environmental Protection Agency (2020, March). National Overview: Facts and Figures on Materials, Wastes and Recycling. Retrieved from <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/national-overview-facts-and-figures-materials>

Chen, M. (2015, June 24). Someone Has to Sort Your Recycling and It’s a Disgusting and Dangerous Job. TheNation. Retrieved from <https://www.thenation.com/article/archive/someone-has-to-sort-your-recycling-and-its-a-disgusting-and-dangerous-job/>