

Capstone Project Proposal



Kamran Taherpour

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	<p>Project: Waste recycling classification</p> <p>This project helps waste management industry to classify garbage into recycling process and reduce manual recycling selection by human.</p> <p>Using computer vision approach and train ML/AI model (CNN) can help to classify waste into different categories like paper, plastic, metal and glass more accurately.</p>
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	<p>Americans throw away 2.5 million plastic bottles every hour, and every day, each person living in the U.S. generates about 4.4 pounds of trash. Ref: https://www.thoughtco.com/waste-disposal-and-recycling-4034910</p> <p>Globally, the annual solid waste is expected to reach 2.2 billion tons by 2025, which would cost \$375.5 billion in waste management Ref: D. Hoornweg and P. Bhada-Tata, What a Waste: A Global Review of Solid Waste Management, World Bank, Washington, DC, USA, 2012.</p> <p>With rapid population growth and urbanization, annual waste generation is expected to increase by 70% from 2016 levels to 3.40 billion tons in 2050. Ref: https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management</p> <p>Proper waste recycling helps environment (reducing</p>

	landfills, water pollution, etc.), public health, and economy (recovering resources and saving energies, reducing cost of manual waste classification).
Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	I'll use AI/ML for image classification of waste to different recycling categories (papers, glass, plastic and metals) and using robot arms to select these objects. This way can reduce environmental problem and landfills, and help economy. It helps also if people confused sometimes to select correct recycle object in the right recycle bin so could lead recycling process longer by manual classification at waste management facility.

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	By gathering statistic about: <ul style="list-style-type: none"> - reducing waste disposal in waterways -reducing landfill for disposal -reducing waste disposal burning at nature - lower cost of manual waste classification Comparison will be stats before and after implementation of waste recycling model.
--	--

Data

Data Acquisition

Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?

We can get open source dataset image and take from actual object too. These data is available without cost. There is no PII or data sensitivity for waste images and issue to overcome. We can acquire a large batch of data to train the model and can be improved by more sample during developing the tool.

Data Source

Consider the size and source of your data; what biases are built into the data and how might the data be improved?

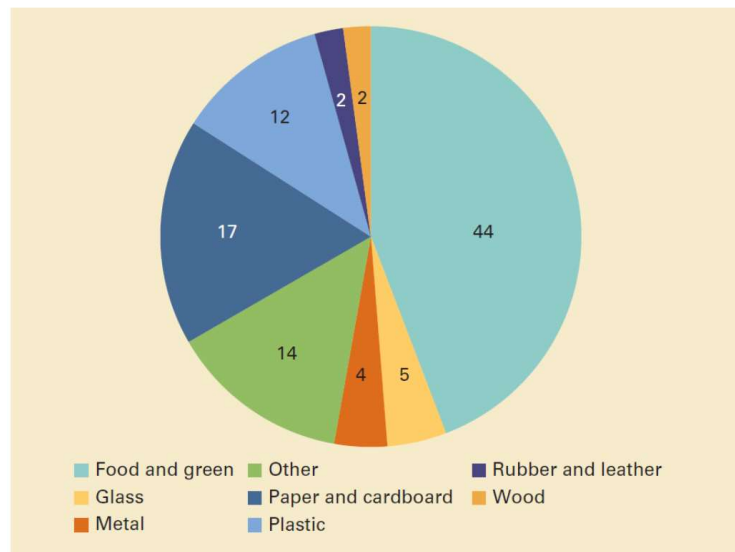
Images from source of data might be taken from different position and angle of object or object might be inside the grey/black bag so difficult to classify. The data can be improved by having more image samples of actual objects in different poses and angels.

Choice of Data Labels

What labels did you decide to add to your data? And why did you decide on these labels versus any other option?

Based on following reference global waste composition, I selected paper, metal, plastic and glass

Global waste composition (percent)



<https://datatopics.worldbank.org/what-a-waste/trends-in-solid-waste-management.html#:~:text=The%20world%20generates%202.01%20billion,from%200.11%20to%204.54%20kilograms>

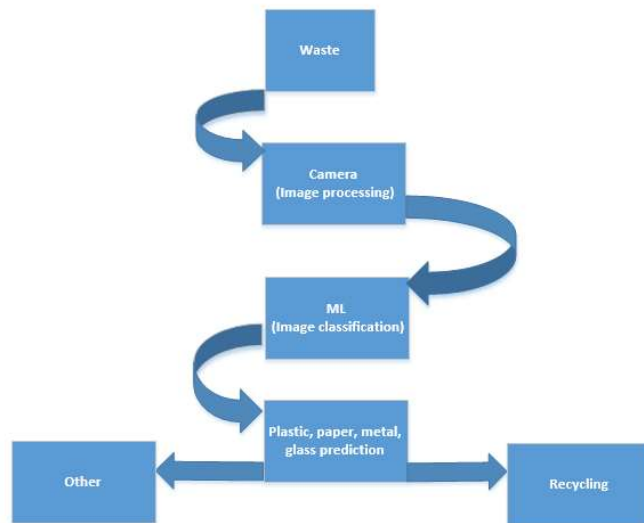
Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	<p>I prefer to use in-house team to build a model after prototype. ML team will do the code for image classification or I can use open-source code out there to use it and team can work with AWS sage maker/Google Auto ML to come up with model, and it can be improved by adding more dataset to have more confidence for good performance.</p> <p>There is separate team will work on robot arm for collection of object too.</p>
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	<p>Since I'm using image classification then precision, recall, confusion matrix and F1-score is good to measure my model success.</p> <p>These are not sensitive data so I would say %85 and above is performance level required for this model.</p>

Minimum Viable Product (MVP)

Design

What does your minimum viable product look like? Include sketches of your product.



Use Cases

What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?

This model is designed for :

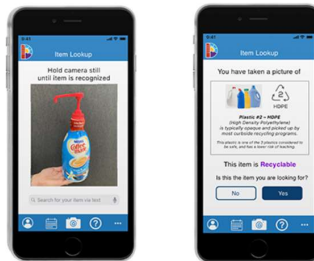
- waste management/recycling facilities

Use cases are recycling waste disposal as much as possible to reduce landfills and others that explained in business cases.

- private sector in waste management industry can get access this product in their facility

The model can be developed to use in public area/airports to educate people for selecting recycling categories before they put in the recycle trash can

This can be developed as an app to educate people at house/public place/airport to select recycling categories also



Roll-out

How will this be adopted? What does the go-to-market plan look like?

- The model should tested as much as we can and resolve issues,

- Meeting with private sectors/Government for demo

- research, marketing and advertisement plan on social media

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	<p>The product will definitely improve in the long-term by having additional data, each object in recycle categories has thousands of samples which leads to new dataset and improve our performance.</p> <p>Real data will be different from training data in terms of shape, size, pose but some data will be the same. The new data can improve precision and recall for our prediction.</p> <p>We can build new model base on new dataset and compare the performance with old one, once we get confidence with new model we can switch (A→B).</p>
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	<p>Team needs regular test and monitor the performance to find unwanted bias during testing before we put into waste management production,</p> <p>Research>> right dataset>>build>>test>>feedback>>fix issues>>lunch</p> <p>Bias is there always but we can reduce by tuning the process.</p> <p>Ref: https://aibusiness.com/document.asp?doc_id=760899</p>