RecycleAI

Materials recycling using machine learning

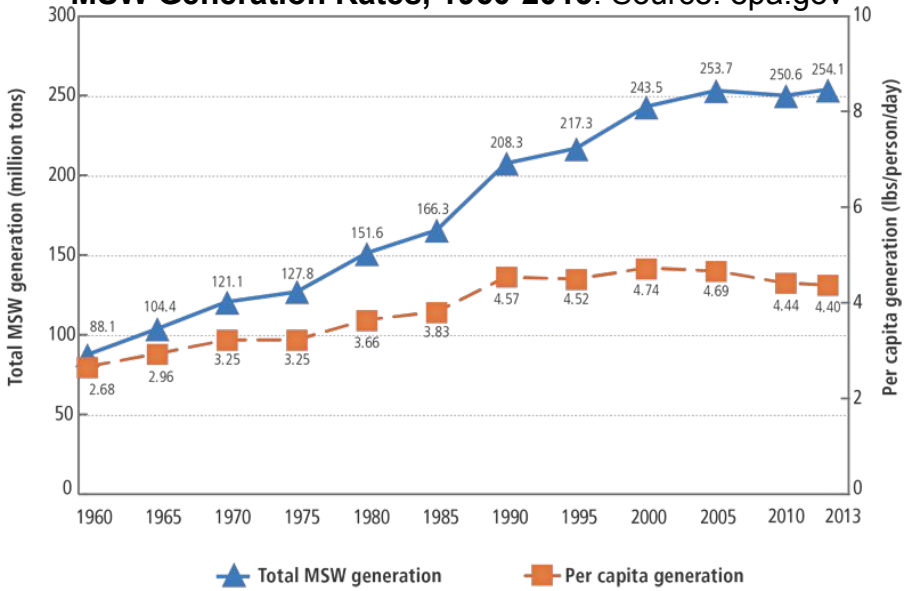
Data-X: Fall 2017

Wesley Jin, Kenny Kang, Elias Orellana, Umesh Thillaivasan, Nate Tran

About **75%** of the U.S. waste stream is **recyclable**, but only **~34**% is recycled

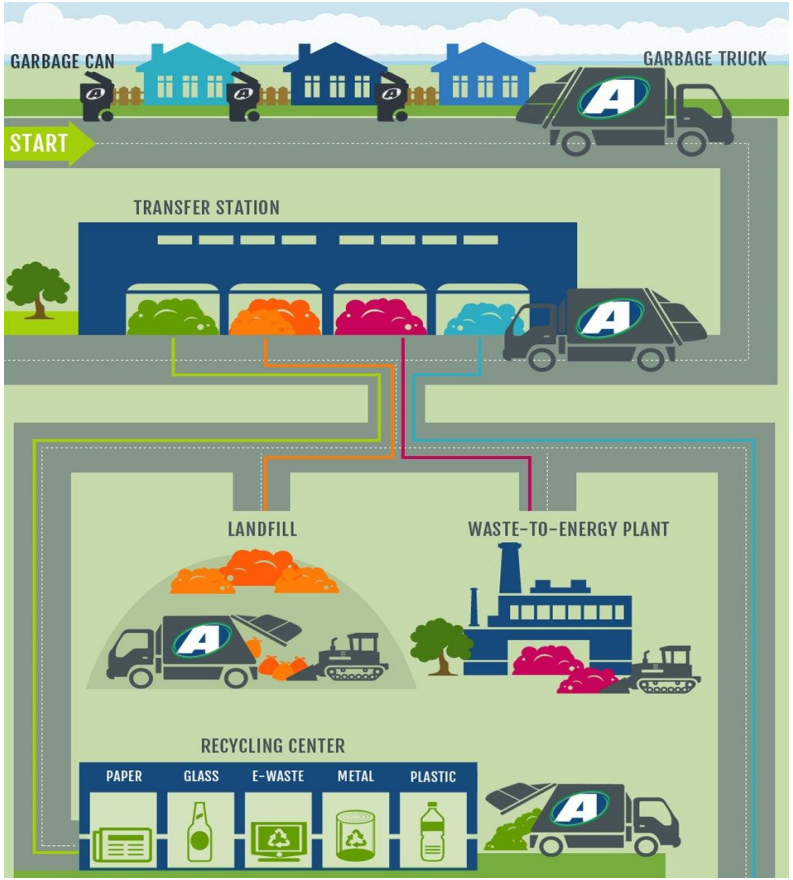
Source: EPA

MSW Generation Rates, 1960-2013. Source: epa.gov



Current Practices

- Only waste in recycling bins ends up at recycling sorting facility
  - Consumer mistakes
  - Sorted by humans and machines
- Trash straight to landfill with no processing



advanceddisposal.com

Current Practices

recycle.georgetown.org

“Single-Stream” Recycling

media.npr.org

Project

Train computer vision system to **classify waste objects.**

- Automate sorting
- Gather and use data for a data-less industry
- Reduce waste going to landfills

Presentation Outline

1. Requirements
2. Architecture
3. Intended UI
4. Data
5. Model Development
6. Results & Demo
7. Challenges
8. Next Steps

Commercialization of RecycleAI

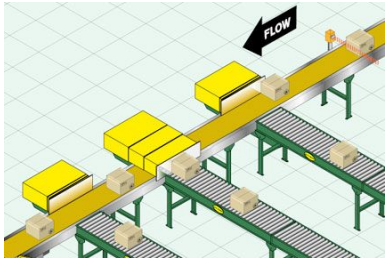
1. Image taken of waste object and input into model

Our Project

2. Model classifies waste object

3. Object sorted to its appropriate destination

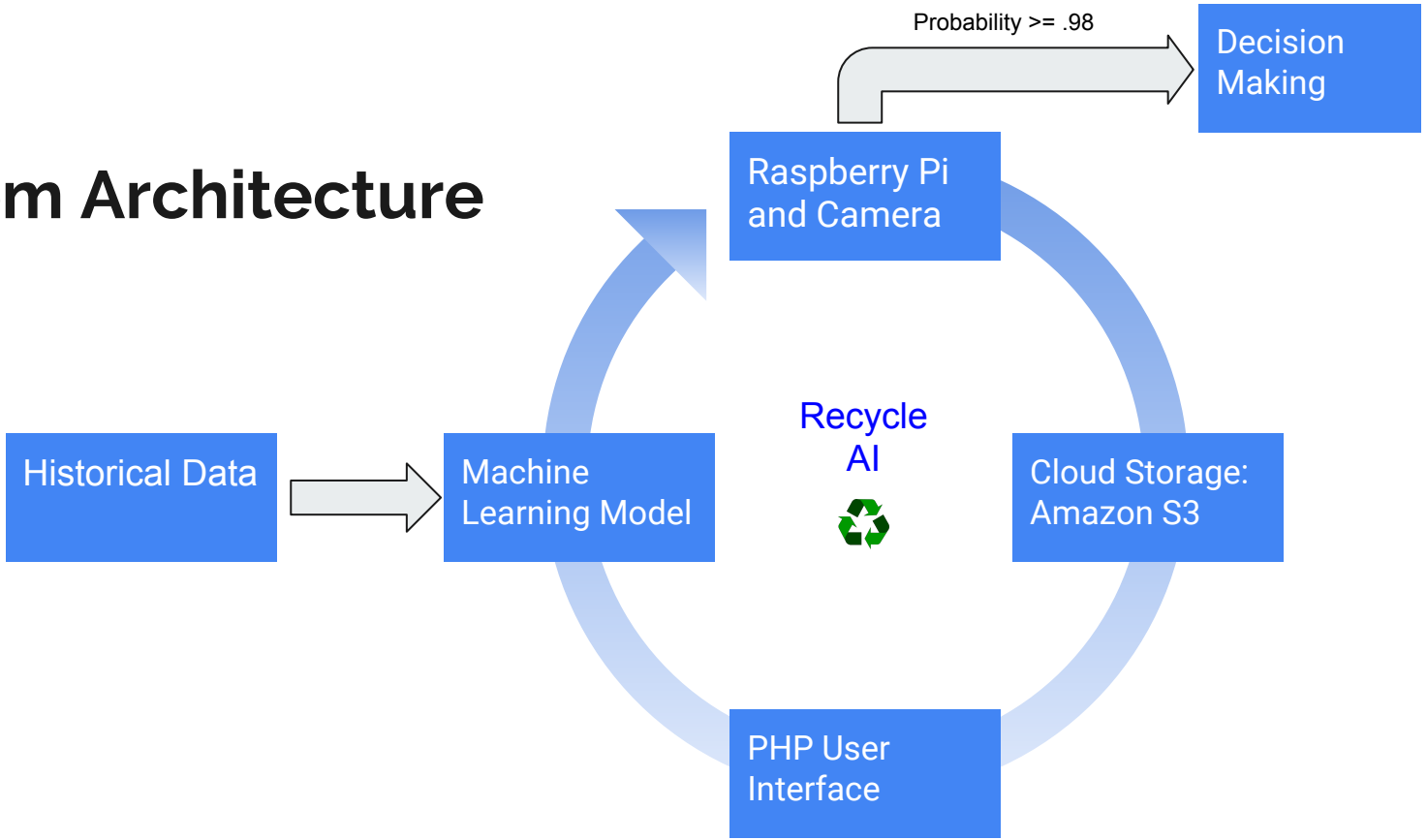
- Bin Sorter
- Robots
- Conveyor Belts

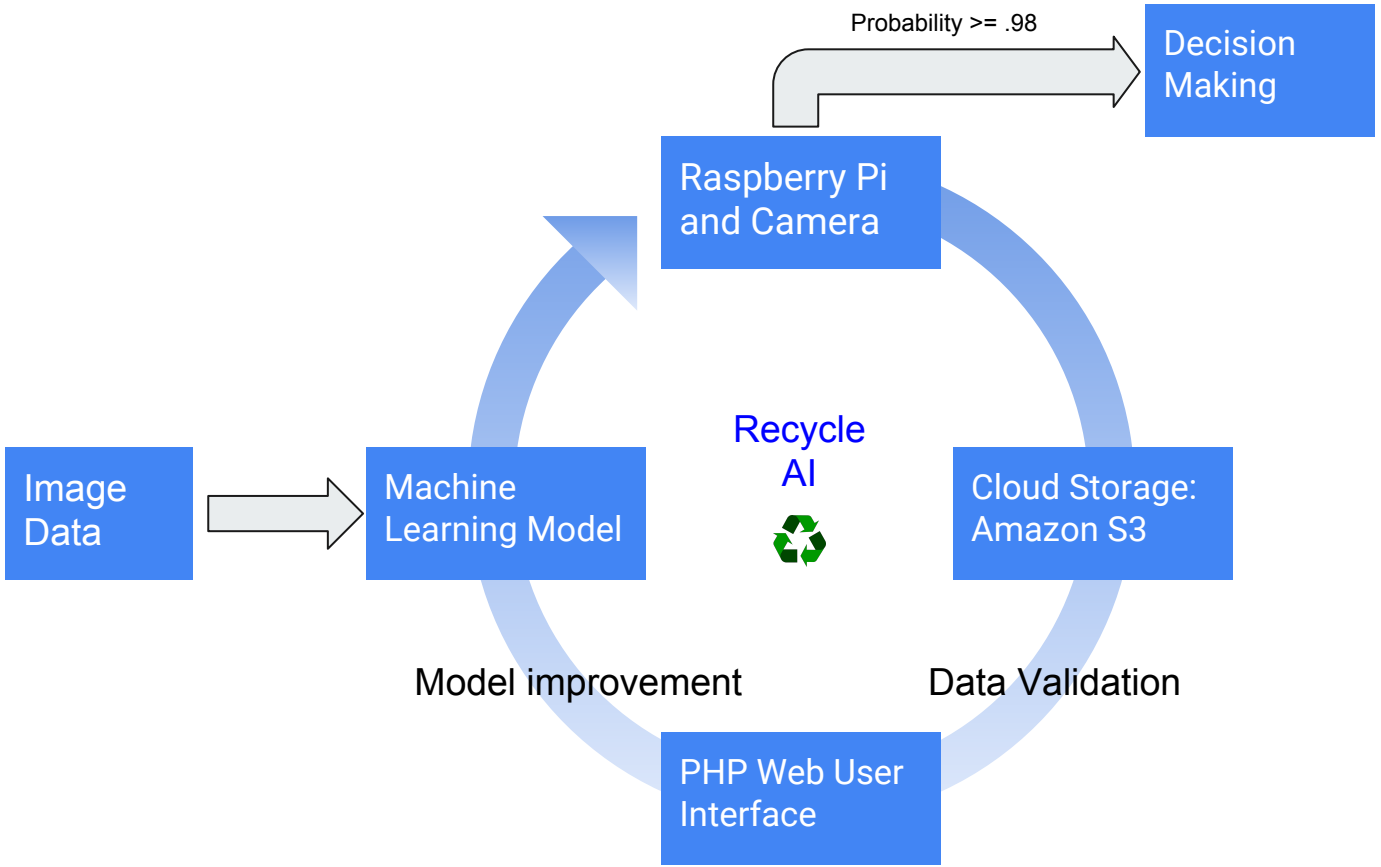


Requirements of Classification System

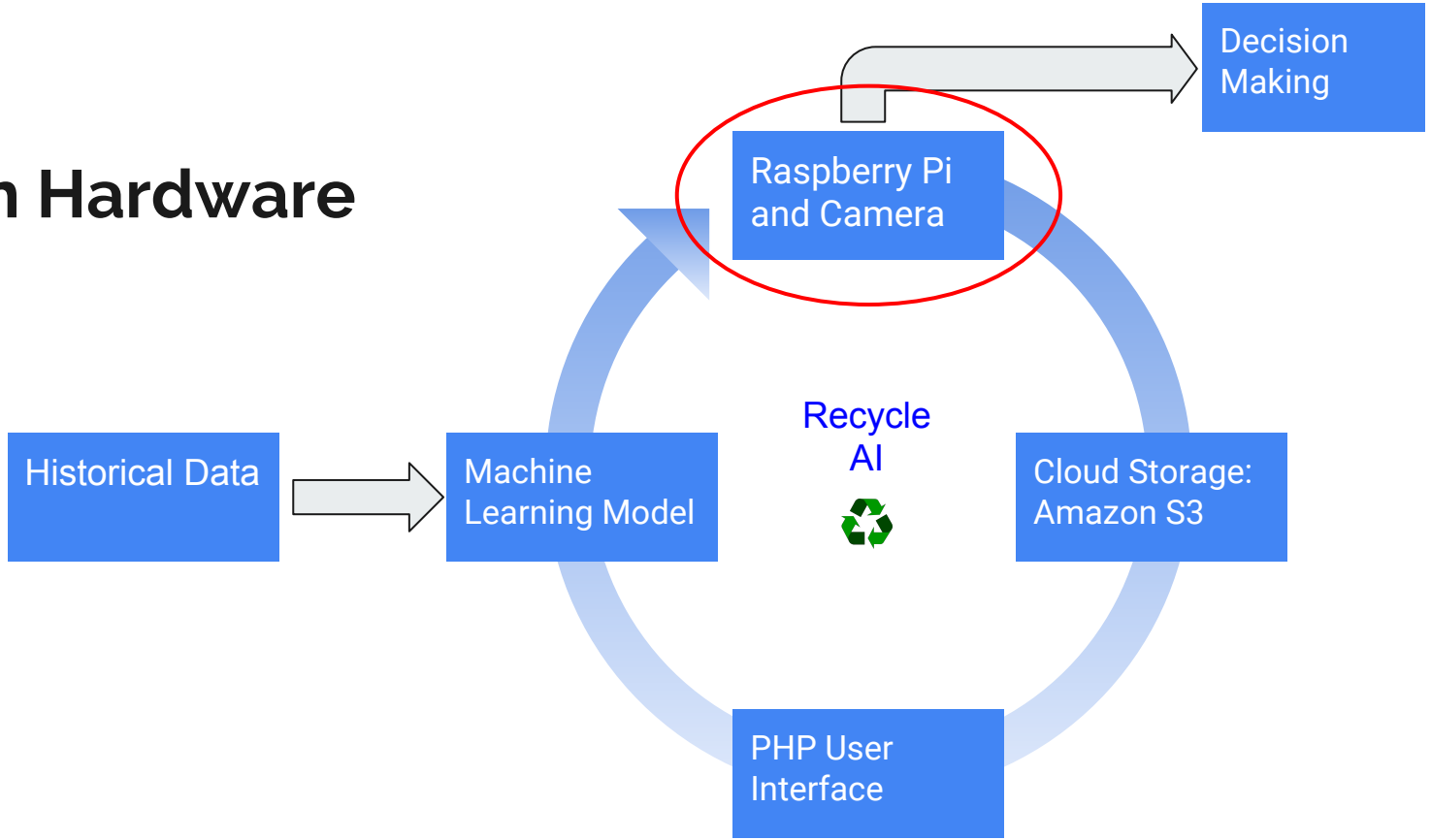
- **Versatile:** Can classify many different types of waste
- **Accurate:** Don't want non-recyclable/compostable objects ending up where they shouldn't be
- **Fast:** Real-world use will involve streams of waste objects
- **Learning/Fine-tuning:** Use mistakes and verified data to make model more robust

System Architecture



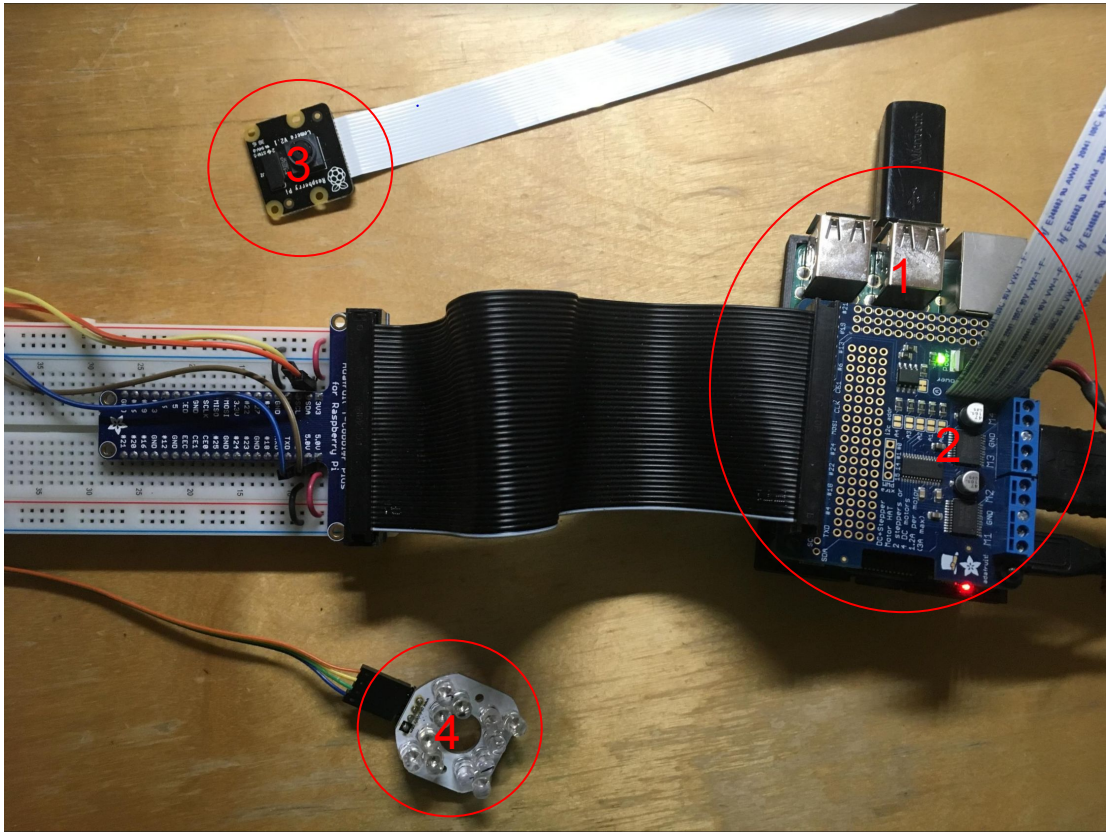

Probability >= .98
Decision Making
Raspberry Pi and Camera
Recycle AI
Image Data
Machine Learning Model
Cloud Storage: Amazon S3
Model improvement
Data Validation
PHP Web User Interface

Vision Hardware

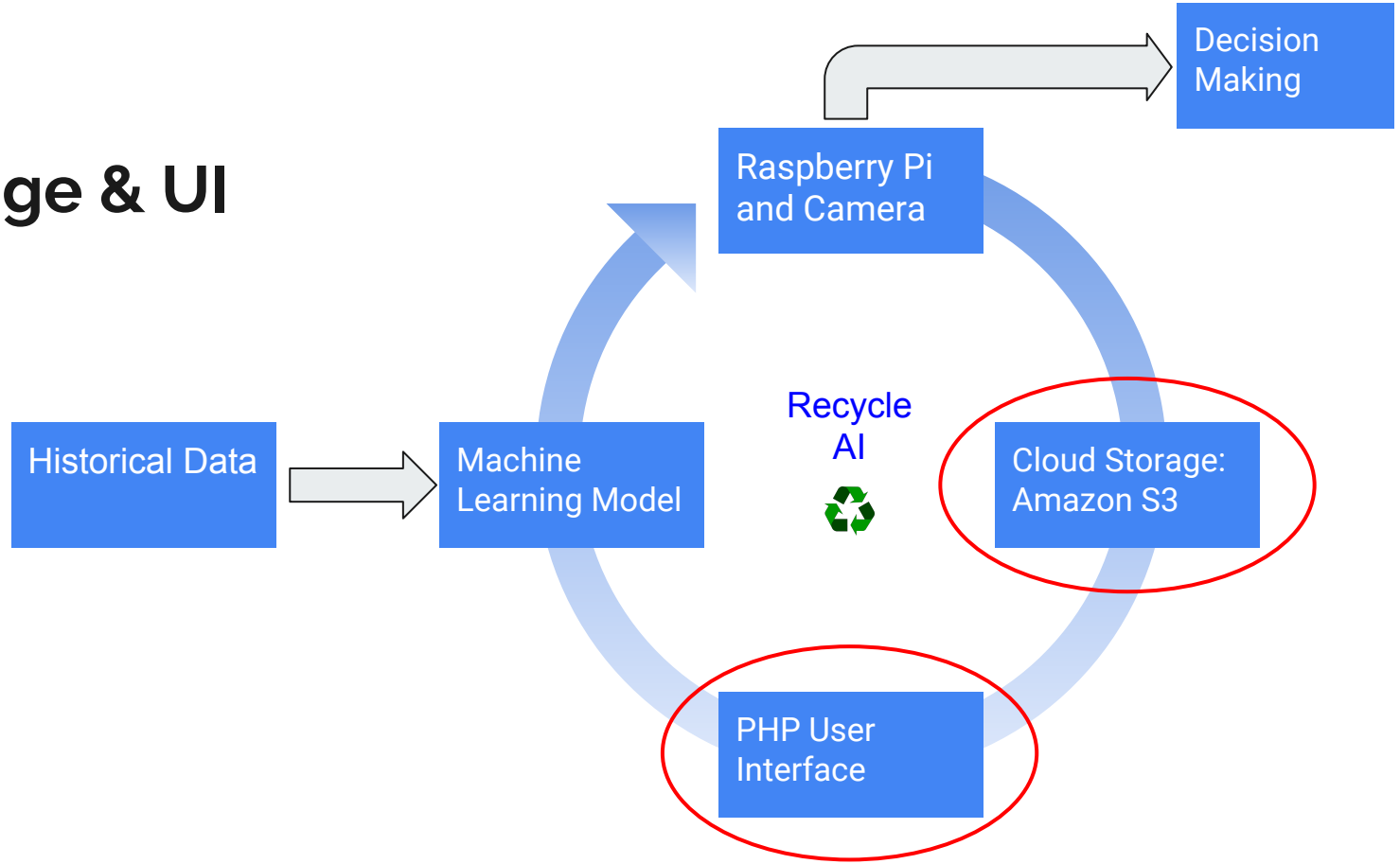


Hardware Components:

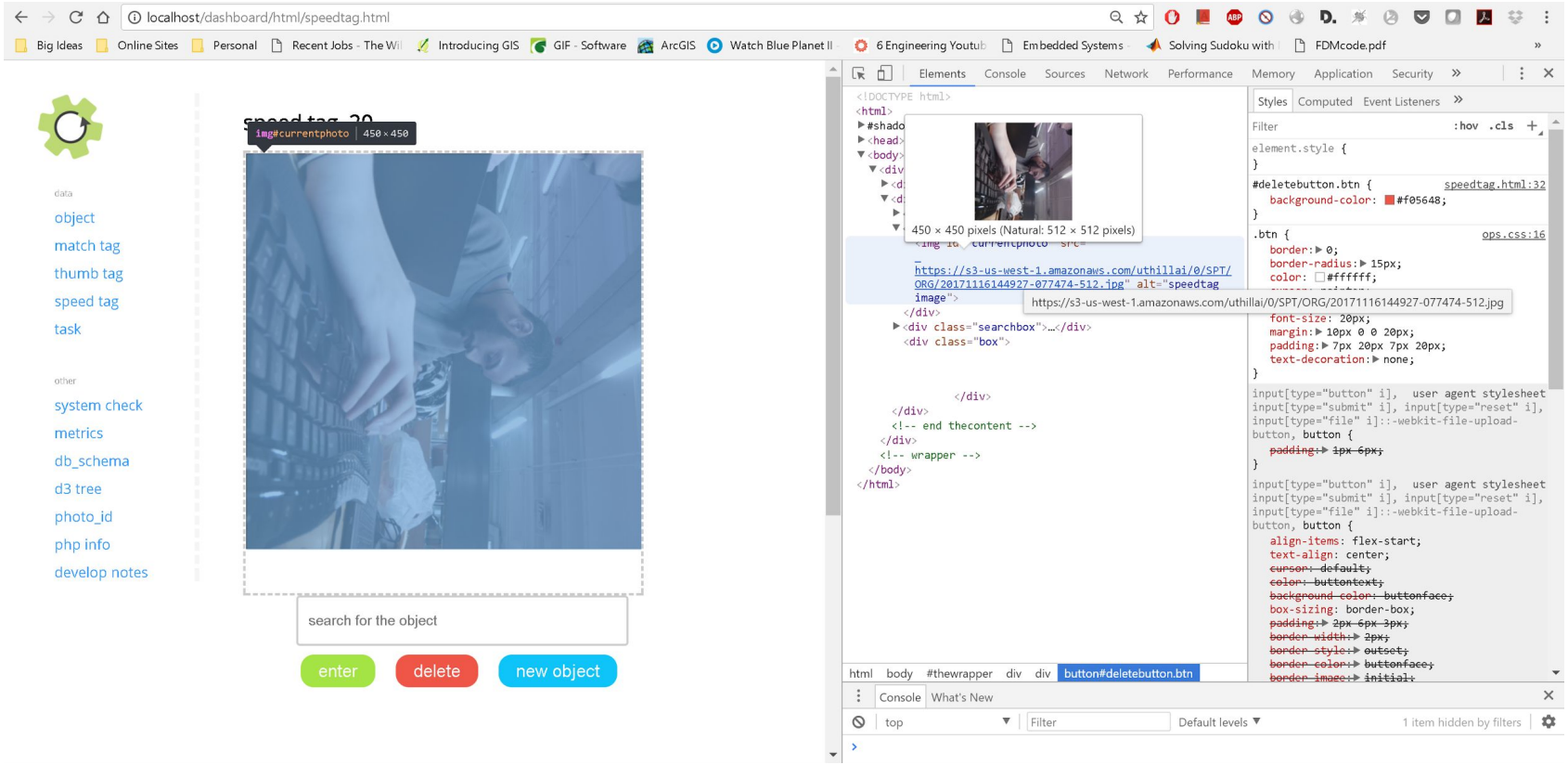
1. Raspberry Pi
2. Circuit board
3. Pi-Camera
4. Bright Pi



Storage & UI


Decision Making
Raspberry Pi and Camera
Historical Data
Machine Learning Model
Recycle AI
Cloud Storage: Amazon S3
PHP User Interface

Intended User Interface



ML Model Development

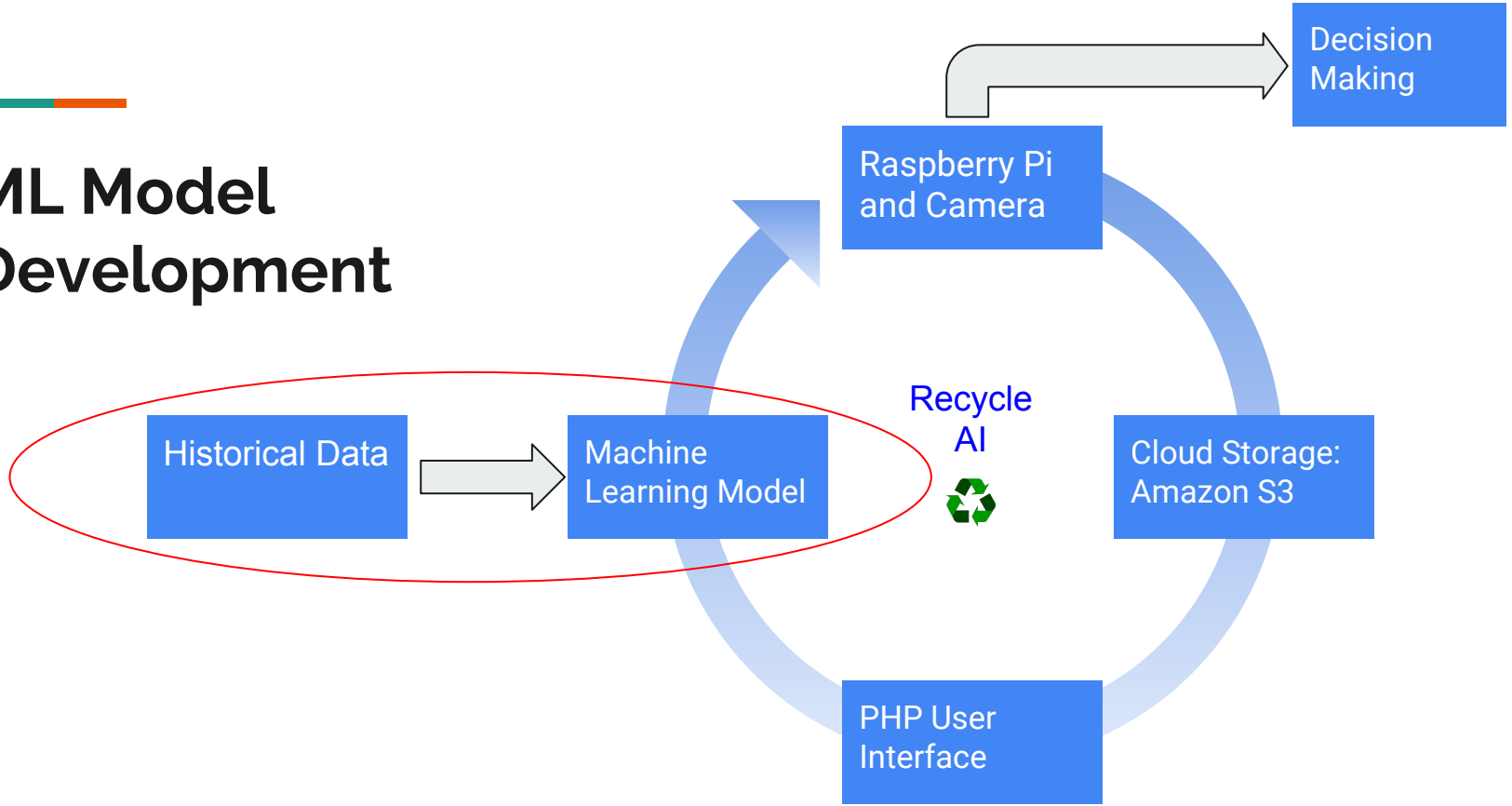


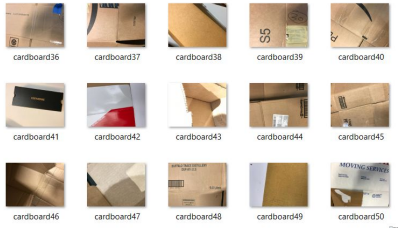
Image Data

75/25
Test/Train Split

183		landfill	plastic film	food wrapper	saran wrap
182		landfill	paper glassine	food wrapper	instant oatmeal
181		compost	organic	fruit	lemon
180		recycle	paper carton	paper	cardboard toilet paper roll
179		recycle	cardboard	paper	tissue kleenex box

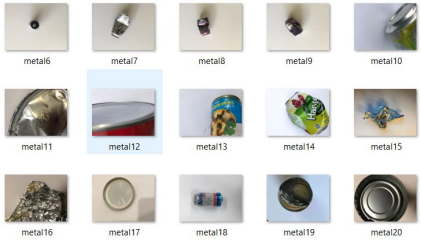
Small (111 objects) existing database created by project mentor Gerry Pesavento - detailed but not comprehensive

Cardboard



Glass

Metal



and images of:

- Plastic
- Paper
- Trash

TrashNet dataset (2527 images) by Gary Thung: github.com/garythung/trashnet

Image Data

Problem: After combining two datasets, still only ~ 3000 images total.

Fix: Augment existing data by flipping/zooming/rotating images




Deep Neural Net Model Development

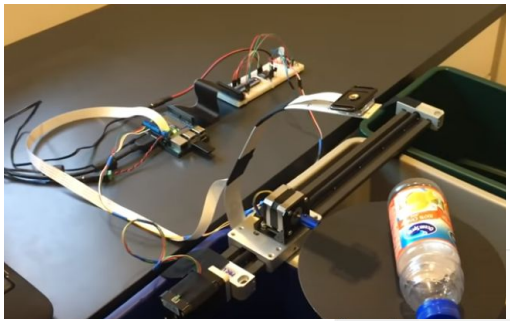
Our approach & learning path

```
model = Sequential()
model.add(Conv2D(32,(3,3), input_shape= input_shape))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2,2)))

model.add(Conv2D(32,(3,3)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2,2)))

model.add(Conv2D(64, (3,3)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2,2)))

model.add(Flatten())
model.add(Dense(100))
model.add(Activation('relu'))
model.add(Dense(100))
model.add(Activation('relu'))
model.add(Dense(6))
model.add(Activation('sigmoid'))

model.compile(loss=keras.losses.categorical_crossentropy,
        optimizer=keras.optimizers.Adadelta(),
        metrics=['accuracy'])

print(model.summary())
```

Approach #1:
Transfer Learning with Fully Connected top layer

Original proposed and used in ‘rAI’ project by Gerry :

- **Originally ~10% Accuracy**
- Difficulty setting up hardware and seeing how all the pieces fit together (Raspberry Pi, PiCam, Linear actuator, AWS, SQL database, outdated code)
- Started with small dataset (only 111 objects classes total), lots of classes

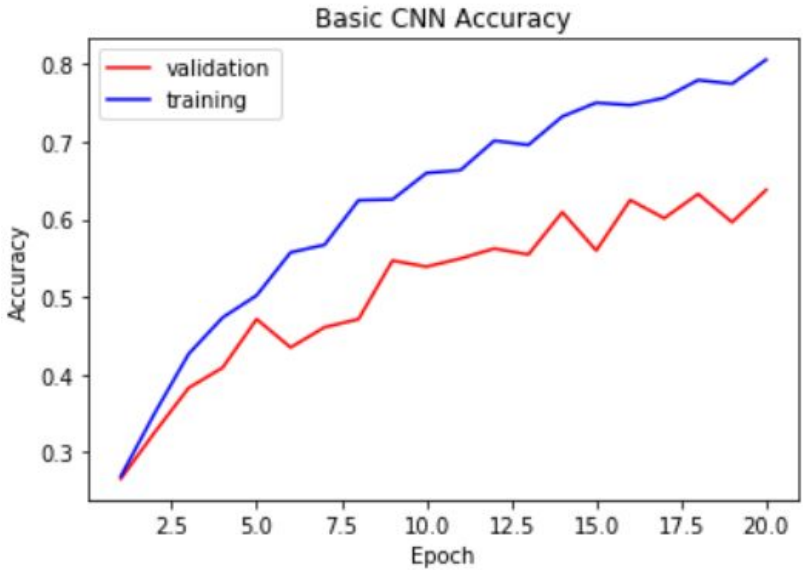


		landfill	plastic film	food wrapper	saran wrap
182		landfill	paper glassine	food wrapper	instant oatmeal
181		compost	organic	fruit	lemon
180		recycle	paper carton	paper	cardboard toilet paper roll
179		recycle	cardboard	paper	tissue kleenex box

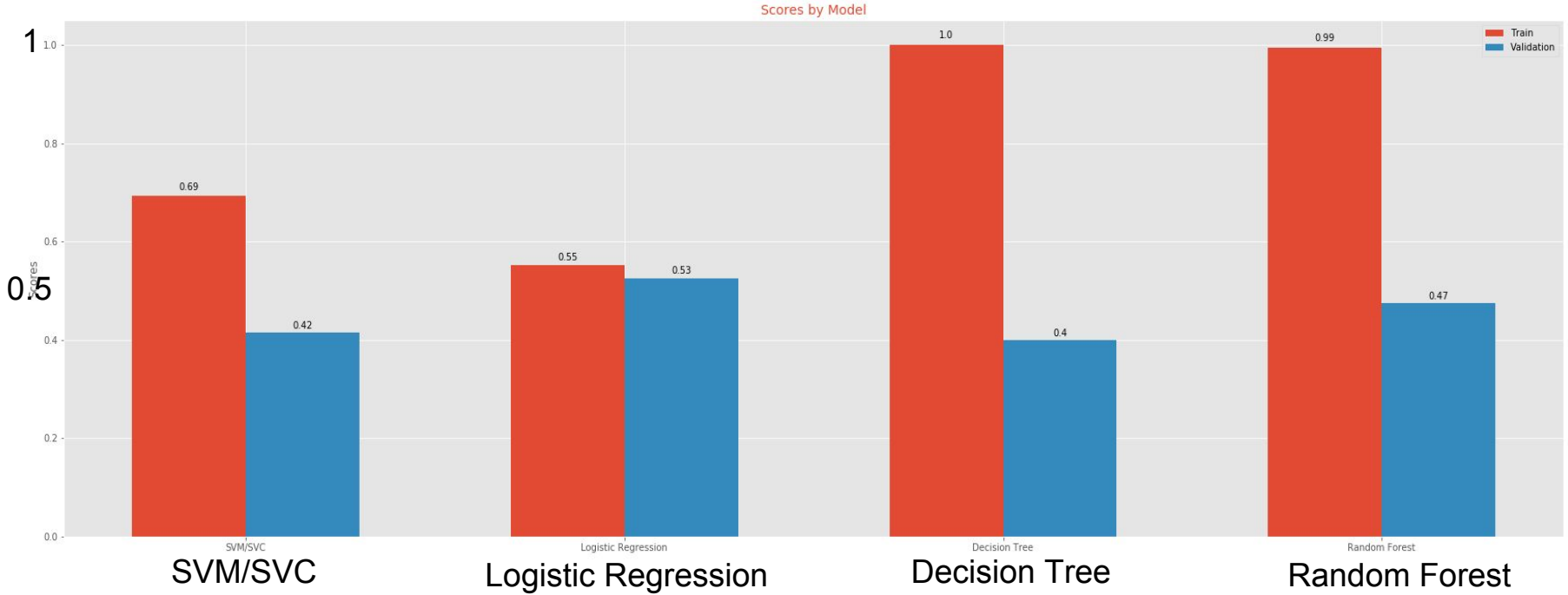
Approach #2: Basic CNN Model from scratch

‘Sequential’ Model as baseline

- **64% Validation Accuracy**
- 3 Convolutional Layers
  - filter sizes: 32, 32, 64
  - 3x3 stride & ReLu activation
- 2 fully connected hidden layers
  - Size 100 & ReLu activation
- Output to 6 categories, Sigmoid activation
- Categorical cross-entropy loss & Stochastic Gradient Descent optimizer
- 20 epochs, 32 Batch size



Approach #3: Transfer learning with basic classification



1

0.5

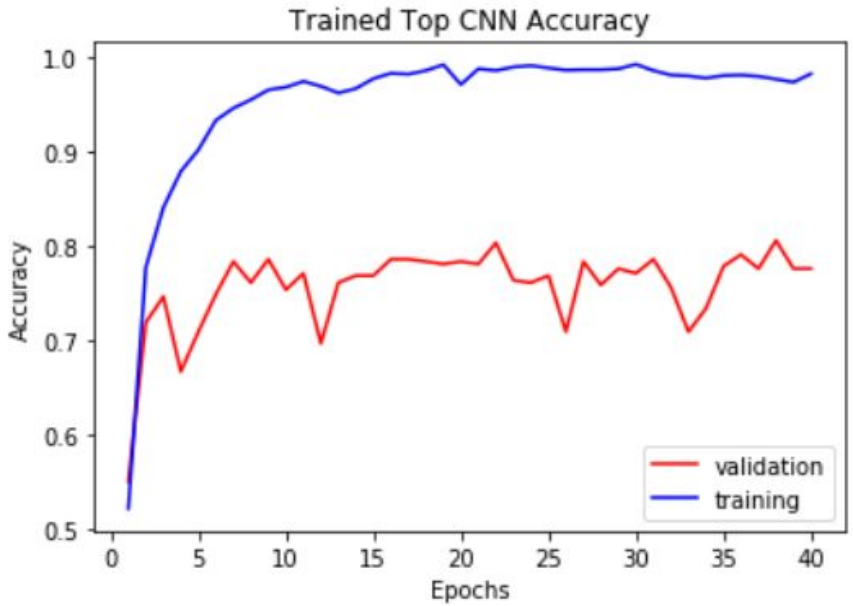
SVM/SVC

Logistic Regression

Decision Tree

Random Forest

Approach #4: Improve Mentor's Model

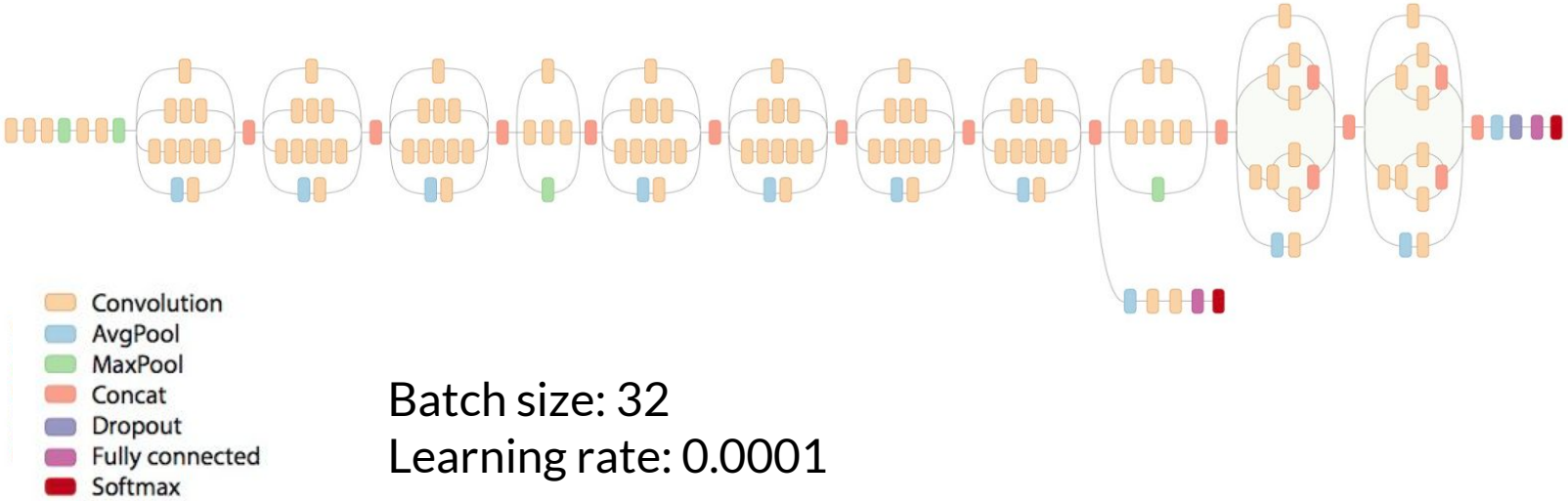


- Only use top layer (high level features)

- **77.6% Validation Accuracy**
- Updated InceptionV3
- Trained on new pre-labeled dataset, with 6 classes

Final Approach: Fine tune pre-trained model

We used the InceptionV3 (ImageNet) model, fine-tuned to our dataset.



Batch size: 32
Learning rate: 0.0001

Final Results

- **86.76% Validation Accuracy**
- InceptionV3 Model + Fully Connected Layer (3 hidden layers with dropout)

- 1) trained top model(fully connected layer) separately on training data
- 2) fine-tuned intermediate weights of the inception model with the same training data.

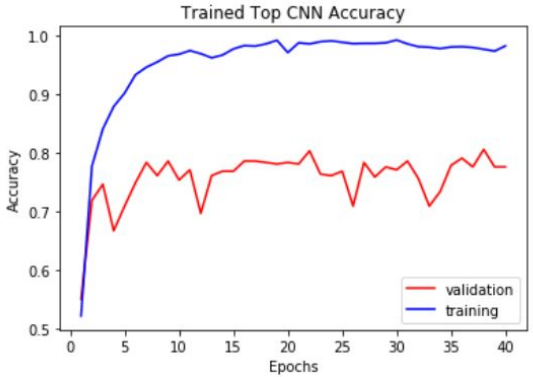
Final Results

SVM:
63% Test Accuracy

Neural Net:
23% Training Accuracy
27% Test Accuracy

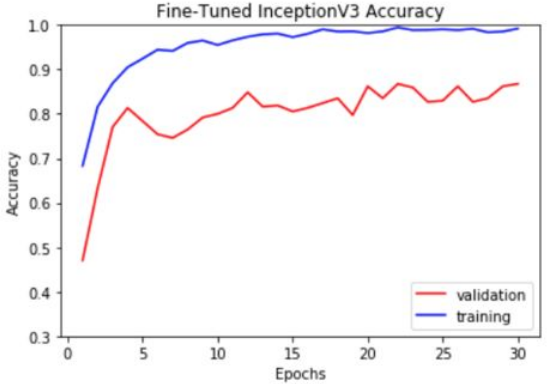
Stanford
Trashnet Model

Source:
http://cs229.stanford.edu/proj2016/poster/ThungYang-ClassificationOfTrashForRecyclabilityStatus-poster.pdf



Mentor's Model

77.6% Accuracy



Our Model

86% Accuracy

Project Demo

Challenges

1. Reviving Gerry's proof-of-concept
    a. Hardware (raspberry pi + pi camera)
    b. Amazon AWS S3 storage for pictures,
    c. SQL database of dataset,
    d. website functionality using PHP,
    e. Training and Prediction models

Challenges

2. Low availability of labelled material waste datasets to train off of.
3. Binary classification of Recycle or Not was not useful
    a. Trained models for 6-class classification for more utility
4. Fooling model with images of similar appearance but different material

Next Steps

1. Improve our neural net models
    a. Collect more data
    b. Better fine-tuning methods
2. More classifications: E-waste, Types of plastic, Hazards, etc.
3. Classify multiple objects at the same time
4. More sensor inputs - not just how object looks in RGB
    a. E.g: Hyperspectral camera

GitHub Repository with Project Code

https://github.com/uthillai/Data-X-Recycle.ai-Project

Questions?