

# Grocery Shopping Recommendation Algorithm

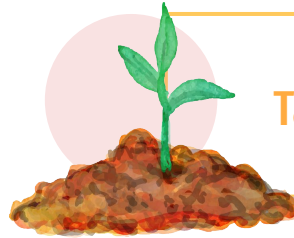
With the food you have and the recipes your prefer,  
for the diet goal you aim to achieve



# Table of Contents

- ➔ **Part I** Project Background
- **Part II** Business and Analytics Plan
- **Part III** Algorithm Development and Performance Evaluation
- **Part IV** Demo
- **Part V** Summary and Potential Next Steps

# Project Objective



## Target Users

- ✓ Desire to eat more **Natural foods**
- ✓ With a **Fast-paced** lifestyle
- ✓ Struggle on **Fitness** recipe choices



## Our Focus

Develop a user-friendly mobile app for:

1. **Accurate food item identification**
2. **Food safety and quality identification**
3. **Customized recipe suggestions**



## Current Fitness and Diet App

- ✗ Personalized grocery shopping lists
- ✗ Options for allergies & intolerances
- ✗ Meal planning with nutritional tracking

# Table of Contents

- **Part I** Project Background
- ➔ **Part II** Business and Analytics Plan
- **Part III** Algorithm Development and Performance Evaluation
- **Part IV** Demo
- **Part V** Summary and Potential Next Steps

# Business Plan



User upload a photo of a natural ingredient



Image classification 1:  
What is this ingredient?

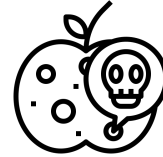
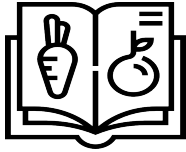
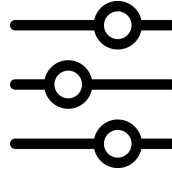


Image classification 2:  
Is it fresh or rotten?



Find  
n recipes

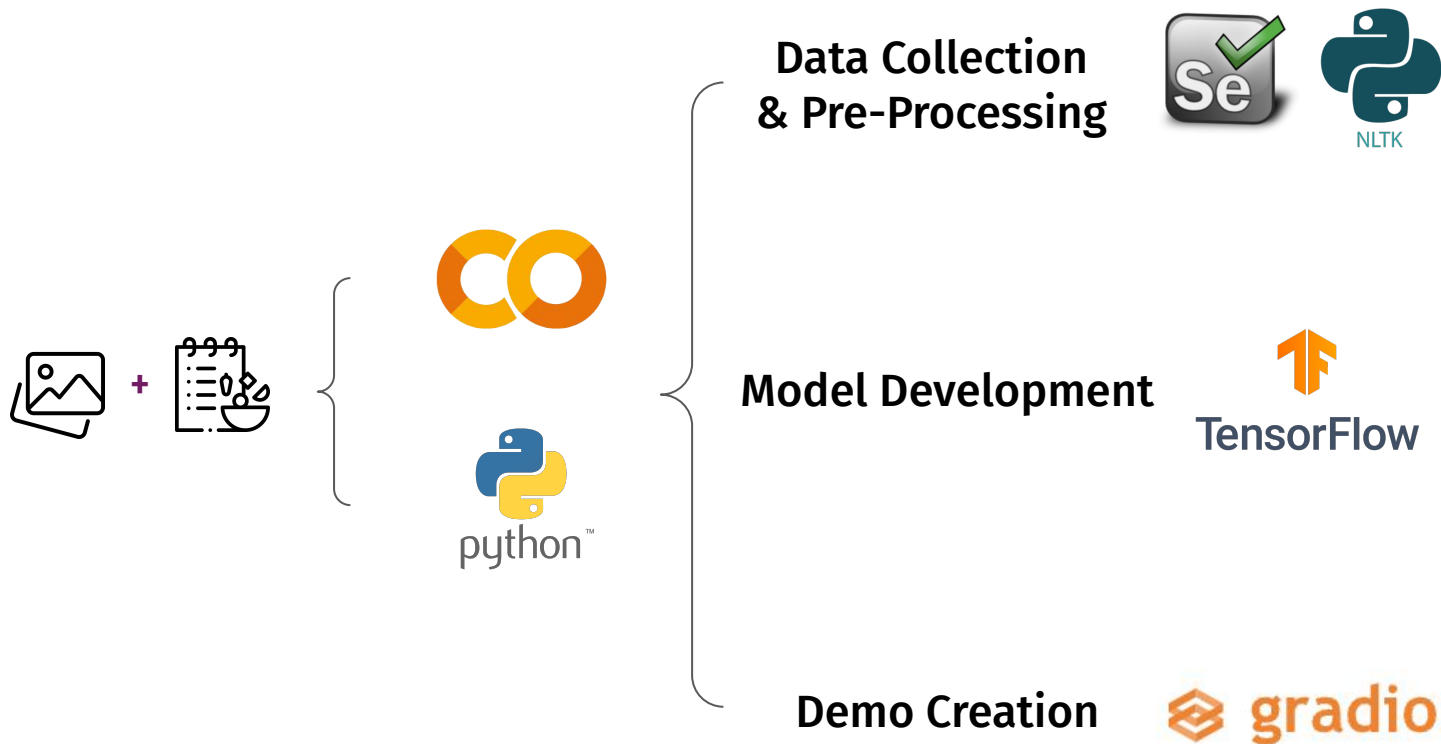


Recipe Filtering



User chooses  
preferred recipes

# Technical Solutions



# Data Source & Analytics Success Metrics

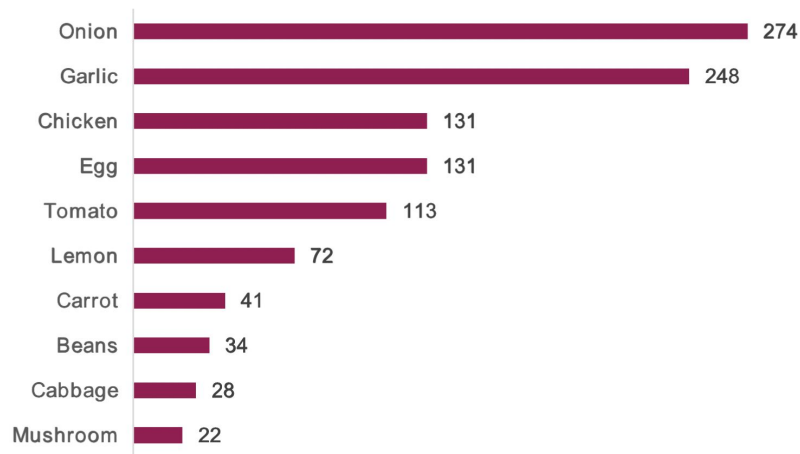
	Classify Image	Recommend Recipes	Customize Shopping List
Data Source	<b>Ingredient Classification</b> V1: 27k <i>scraped</i> pics V2: 27k <i>handpicked+augmented</i> pics V3: a backup set from kaggle	<b>457 recipes</b> scraped from allrecipes.com	<b>3 diet types:</b> Veggie, High Protein, Normal <b>7 styles:</b> CHN, IND, KOR, JPN, GRC, MEX, ITA
	<b>Fresh/Rotten Classification</b> 1 Image Set from Kaggle		
Success Metrics	Classification Accuracy	Number of Recipes Available per Ingredient	Number of Customized Combinations Available (e.g. Vegetarian x Mexican, High Protein x Korean, etc.)

# Additional Information on the Image Datasets



## Choice of Ingredients

■ Number of Recipes



Prioritized ingredients  
that appeared in more recipes.



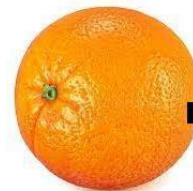
## Image Augmentation Methods



Original



Flip



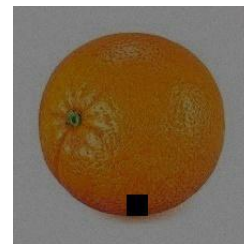
Delete square



Adjust Brightness



Add Noise



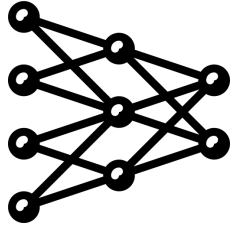
A+B+C+D



# Table of Contents

- **Part I** Project Background
- **Part II** Business and Analytics Plan
- ➔ **Part III** Model Development and Evaluation
- **Part IV** Demo
- **Part V** Summary and Potential Next Steps

# Model Development



## Model Architecture

A CNN neural network with:

- Keras to pre-process image data
- 9 layers to extract features of the images
- class\_mode parameter is set to 'categorical'

## Model Training

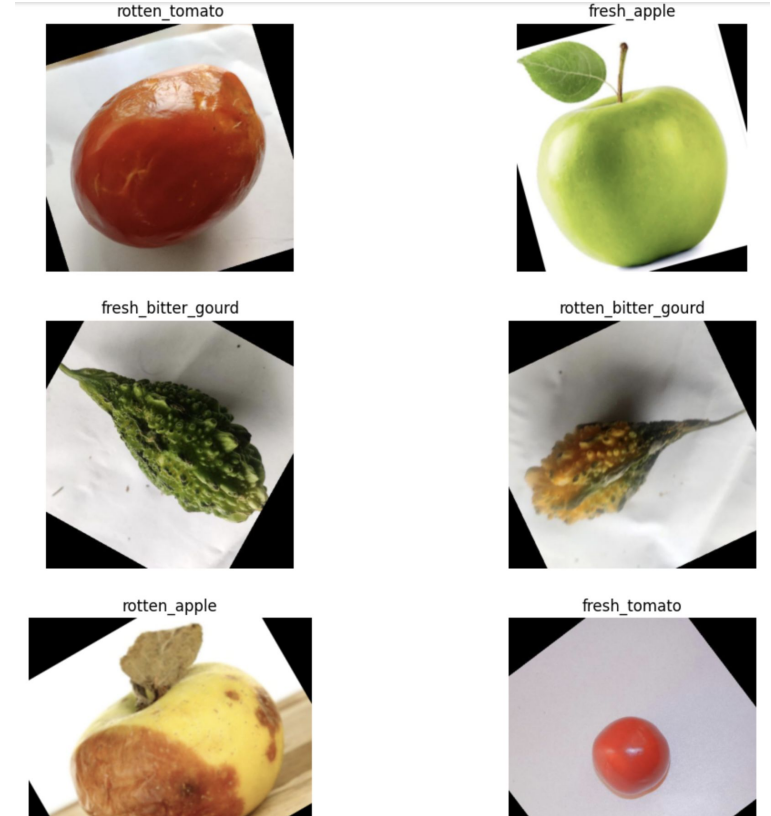
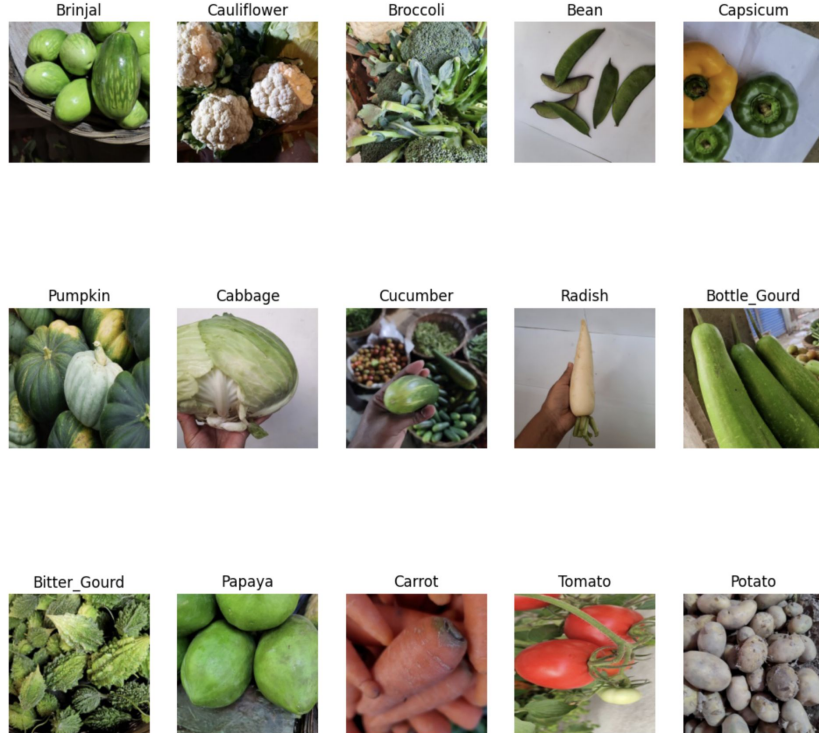
Used the four datasets to train two models respectively

- 65% train, 17.5% validation, 17.5% test

## Model Evaluation

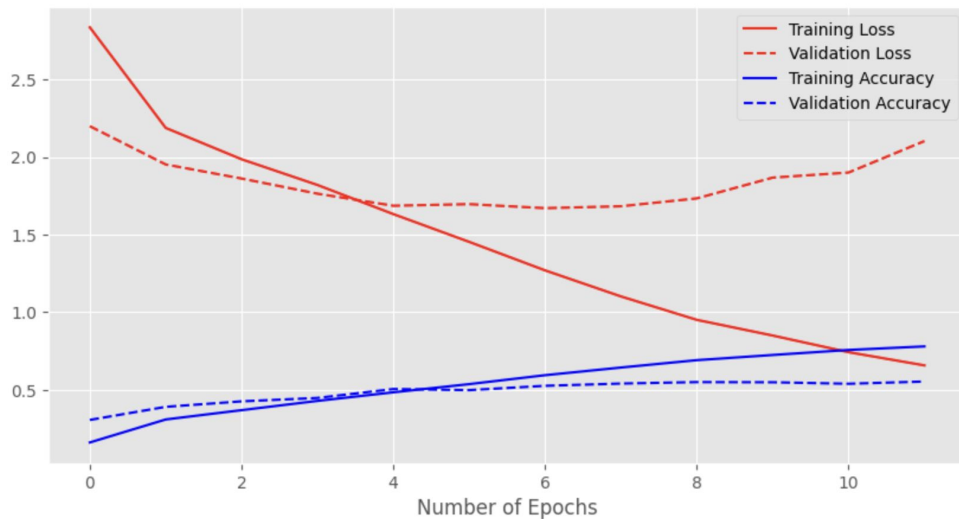
Refer to classification accuracy as model evaluation metrics

# Image Example



# Model Evaluation - Ingredient Classification (1/3)

**Model 1**  
With scraped images

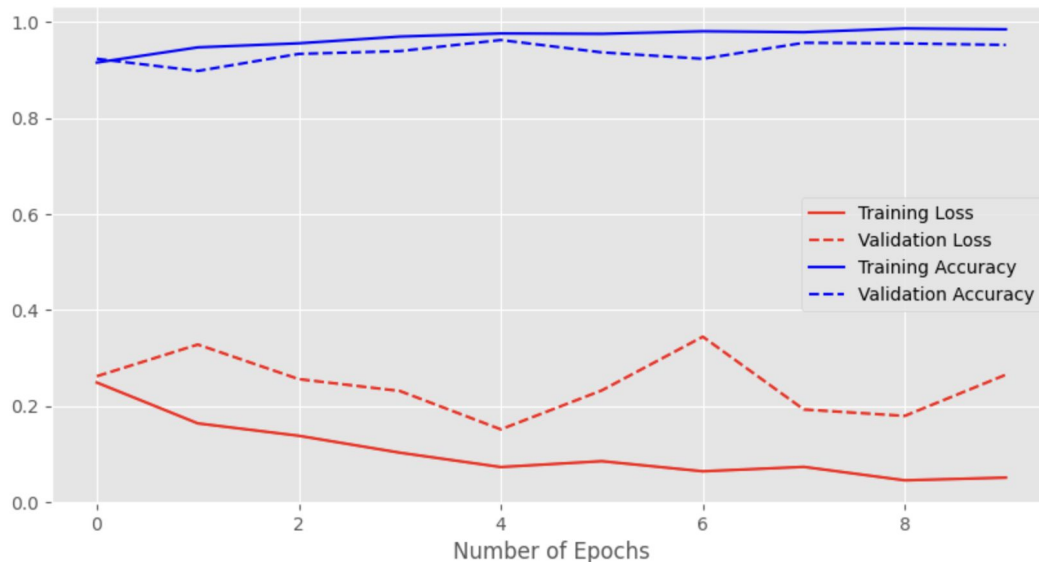


**Evaluation:**  
loss: 1.9843, accuracy: 0.5544

# Model Evaluation - Ingredient Classification (2/3)

## Model 2

With handpicked and  
augmented images



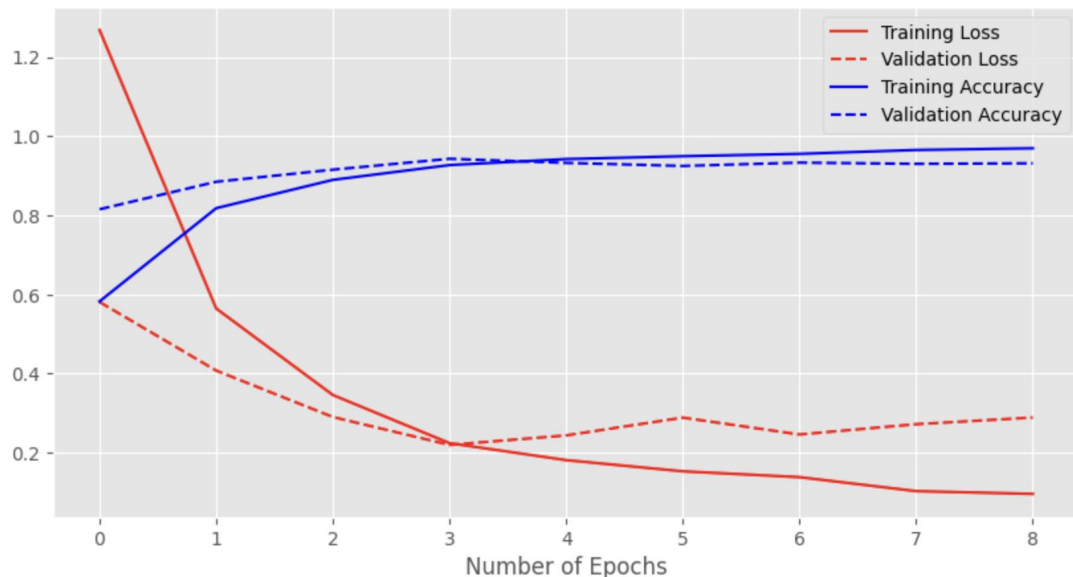
**Evaluation:**

loss: 0.2605, accuracy: 0.9568

# Model Evaluation - Ingredient Classification

## Model 3

With the backup image  
dataset

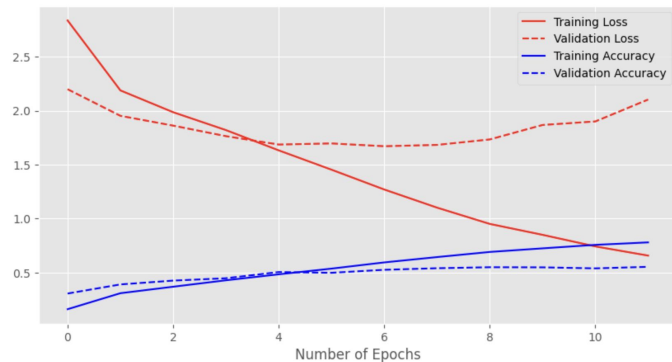


### Evaluation:

loss: 0.2669, accuracy: 0.9337

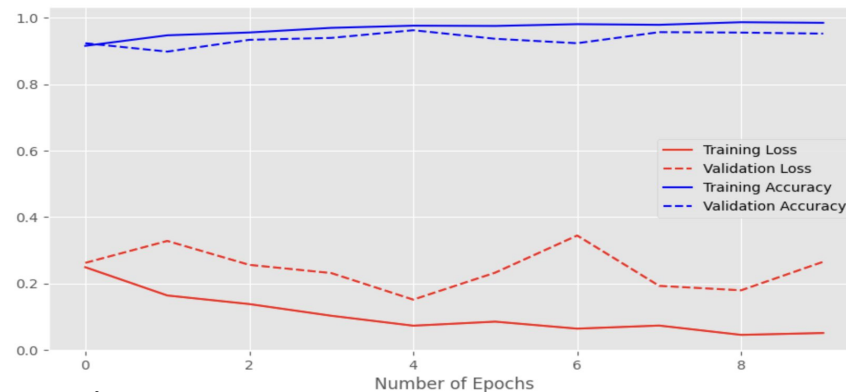
### Model 1 With scraped images

**Evaluation:** loss: 1.9843, accuracy: 0.5544



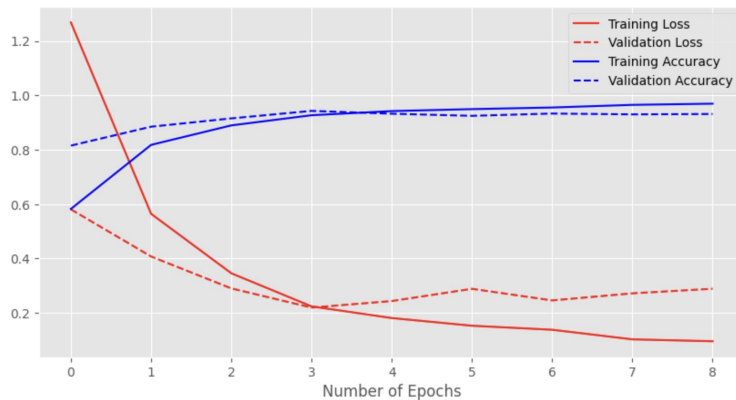
### Model 2 With handpicked and augmented images

**Evaluation:** loss: 0.2605, accuracy: 0.9568



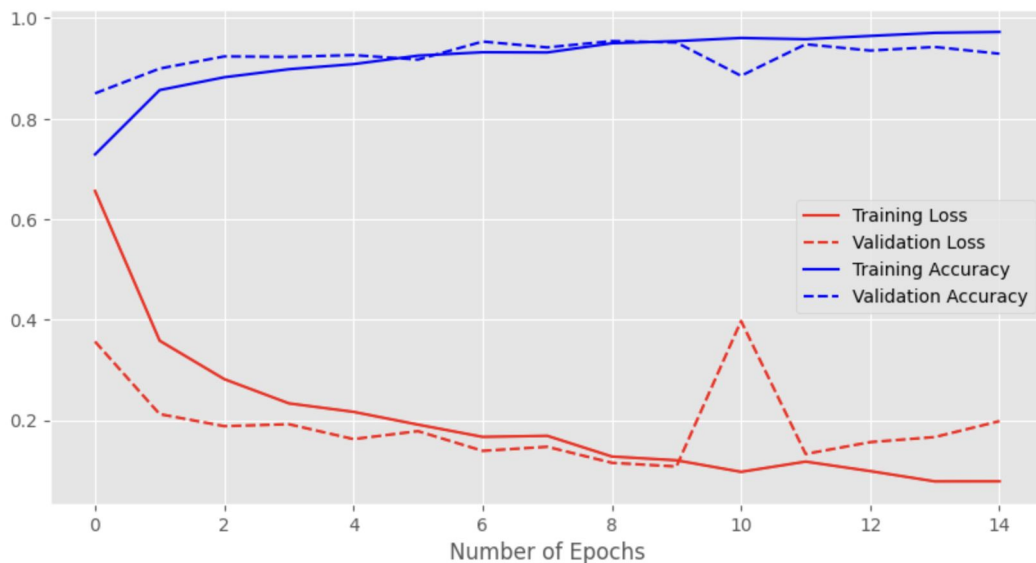
### Model 3 With the backup image dataset

**Evaluation:** loss: 0.2669, accuracy: 0.9337



# Model Evaluation - Fresh / Rotten Classification

**Model 1**  
With Kaggle Dataset



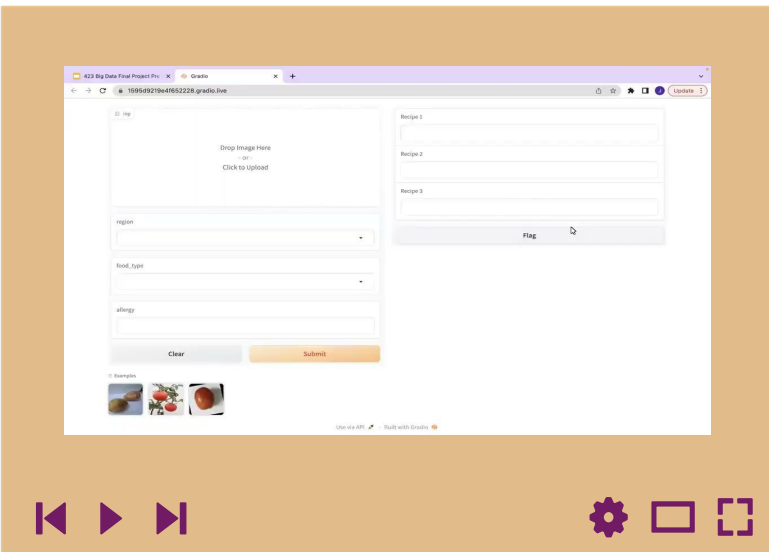
**Evaluation:**

loss: 0.1642, accuracy: 0.9391



# Table of Contents

- **Part I** Project Background
- **Part II** Business and Analytics Plan
- **Part III** Algorithm Development and Performance Evaluation
- ➔ **Part IV** Demo
- **Part V** Summary and Potential Next Steps



# Table of Contents

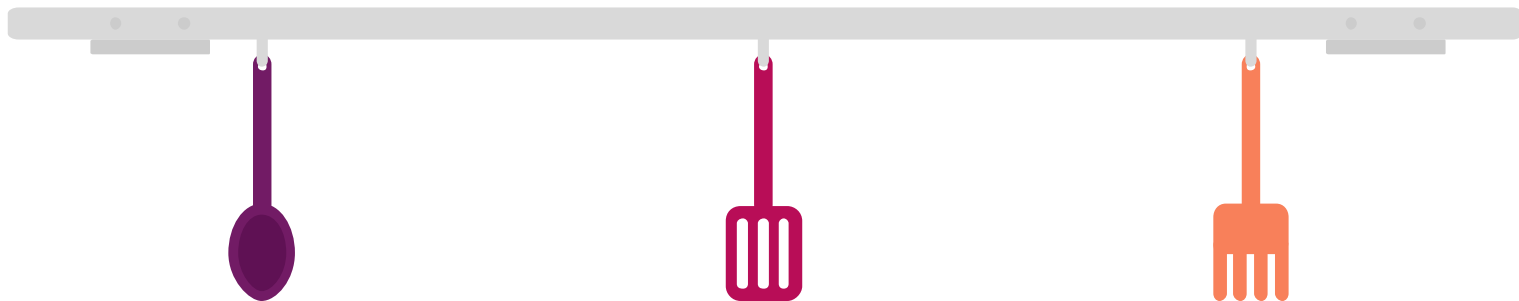
- **Part I** Project Background
- **Part II** Business and Analytics Plan
- **Part III** Algorithm Development and Performance Evaluation
- **Part IV** Demo
- ➔ **Part V** Summary and Potential Next Steps

# Progress Summary

## 1. Originally planned features that have been realized:

- Ingredient classification
- Fresh/Rotten classification
- Customized recipe recommendation

## 2. Successfully built a image dataset with relatively higher classification accuracy



# Reflection and Potential Next Steps

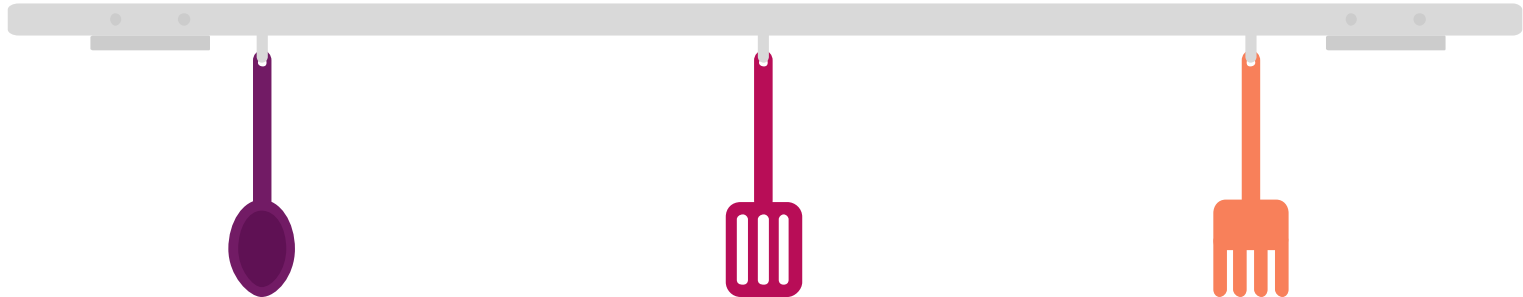
## Reflection

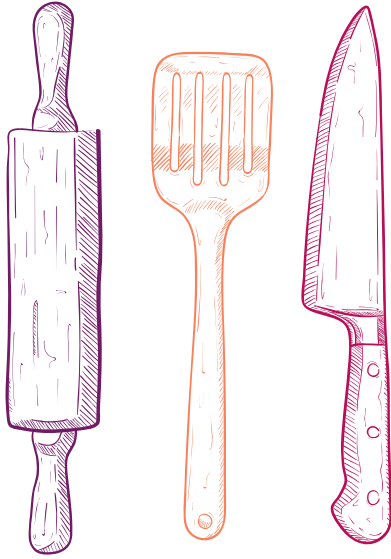
What we could improve:

1. Increase ingredients included in the dataset (currently the number is 30)
2. Increase the number of recipe per ingredient
3. Increase model accuracy and reduce overfitting
4. Improve the bridging of two classification algorithms

## Potential Next Steps

1. Experiment with publicly available image dataset: ImageNet, CIFAR, or Open Images
2. Model tuning: activation function, optimization algorithm, etc.
3. Improve the demo





**Thank you!**