

# DS Final Project Proposal

## 依據天氣影響之下周菜價預測平台

Sorry my bad

- 曾柏誠 E24126717
- 何寬羿 C34104032
- 林業誠 E24105038

# Contents

1. Motivations
2. Expected Dataset to be Crawled
3. Problem Statement
4. Technical Challenges
5. Preliminary Methods
6. Evaluation Plans
7. Expected Time Schedule

# Motivations

菜價破**60**元大關！貴7成「創秋冬新高」 農糧署：颱風導致

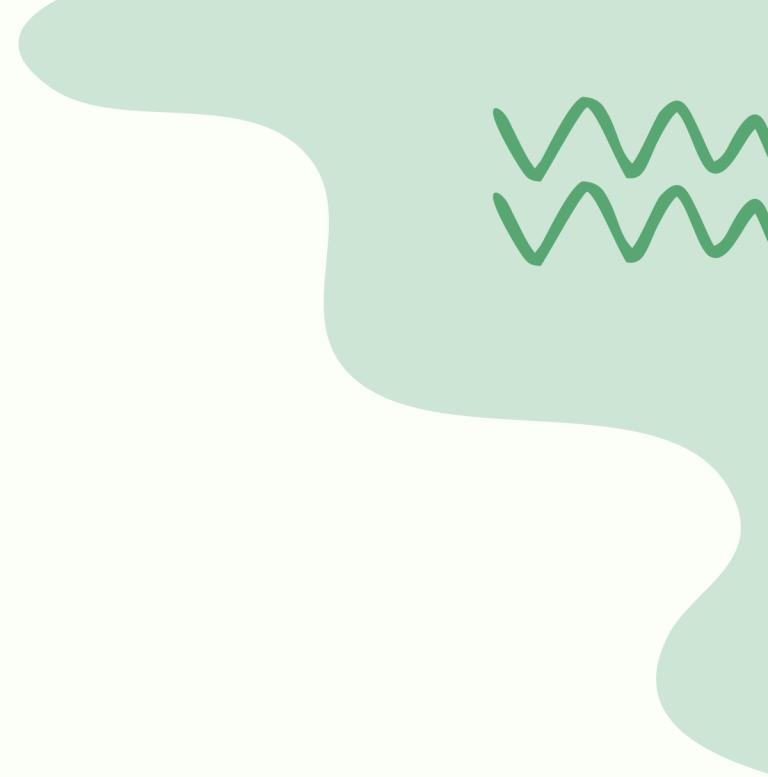
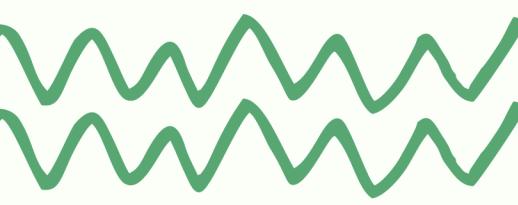


蕭意淮

2024年11月25日 週一 下午5:14



菜價破 60 元大關。（圖 / TVBS 資料畫面）



# Motivations

蔬菜價格受天氣影響極大，導致價格波動。

## 應用

需求者：規劃採購時間和食材選擇

供應者：優化收成及銷售計劃



# Expected Dataset to be Crawled

## 1. 天氣資料 (交通部中央氣象署)

- 特點：氣象資料豐富，更新頻率高、即時。
- 應用：捕捉天氣對農作物生長和供應的影響。

## 2. 菜價資料 (農業部-田邊好幫手)

- 特點：每日更新不同地區的蔬菜價格
- 應用：結合歷史菜價與天氣特徵進行價格預測

## 3. 颱風資料 (颱風資料庫)

- 特點：導致菜價急劇波動的重要因素
- 應用：增強模型預測的準確性

# Problem Statement

## **Input:**

近期天氣資料、近期颱風資料、時間、菜的種類、(種植地點)

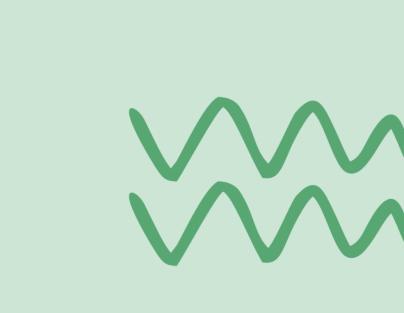
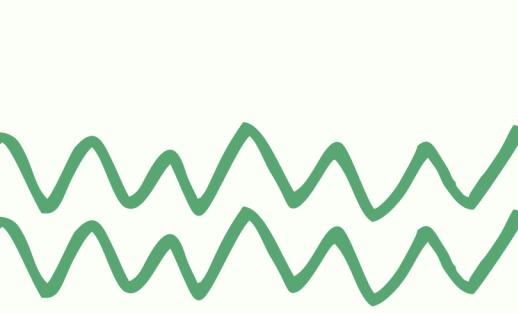
## **Output:**

預測下周某菜類的平均價格

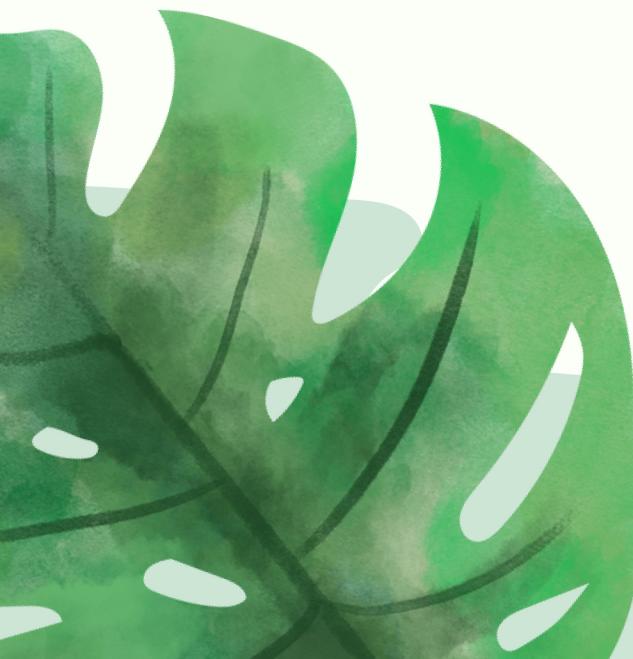
## **※現有問題：目前缺乏預測菜價的平台**

同：多數模型只考慮歷史菜價、單一菜種、缺乏平台

異：納入天氣及颱風資料、(多種菜種)、GUI



# Technical Challenges

1. 影響菜價的因素不只有天氣(政策、市場供需、通膨...)  
=> 準確性
  
  2. 預測多種菜類  
=> 複雜、耗時
- 

# Preliminary Methods

**特徵：**

天氣、颱風、時間、菜種

**模型：**

XGBoost、LSTM等模型

**創新：**

結合天氣數據、不同地區&多種蔬菜、方便使用(GUI)

# Evaluation Plans

## Training Settings:

資料劃分(7:3)、超參數調整、模型訓練

## Evaluation Metrics :

均方根誤差 (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

## Baselines:

傳統簡易機器學習模型

# Expected Time Schedule

W13~14

資料收集&爬蟲設計

W15~16

模型訓練&預測實驗

W17~18

整理報告&成果發表

Thank You