

# การวิเคราะห์และเปรียบเทียบประสิทธิภาพของโมเดลในการทำนาย

## วัตถุประสงค์

เพื่อทำนายค่าของจำนวน Apple จากทั้ง 3 สายพันธุ์ในช่วงเวลาที่กำหนดผ่านการเรียนรู้ด้วยโมเดลที่หลากหลายและเปรียบเทียบผลลัพธ์ที่ได้ว่าโมเดลใดมีประสิทธิภาพดีที่สุดในการทำนาย

## ชุดข้อมูลที่ใช้

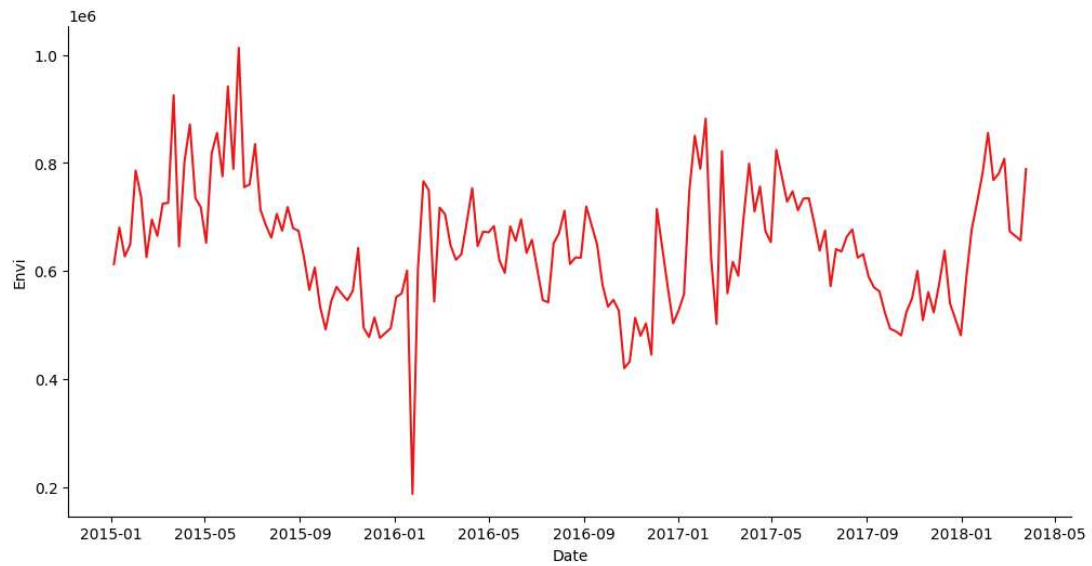
ใช้ชุดข้อมูลจากไฟล์ Apple462.csv โดยมีข้อมูลของ Apple 3 สายพันธุ์ คือ Envi, Gala และ Fuji และ Date เป็นวันที่เก็บจำนวนผล Apple ทั้ง 3 สายพันธุ์ โดยเลือกการทำนายผลจากประเทศ Africa ซึ่งมีจำนวนตัวอย่างทั้งหมด 169 แถว

## การเตรียมข้อมูล

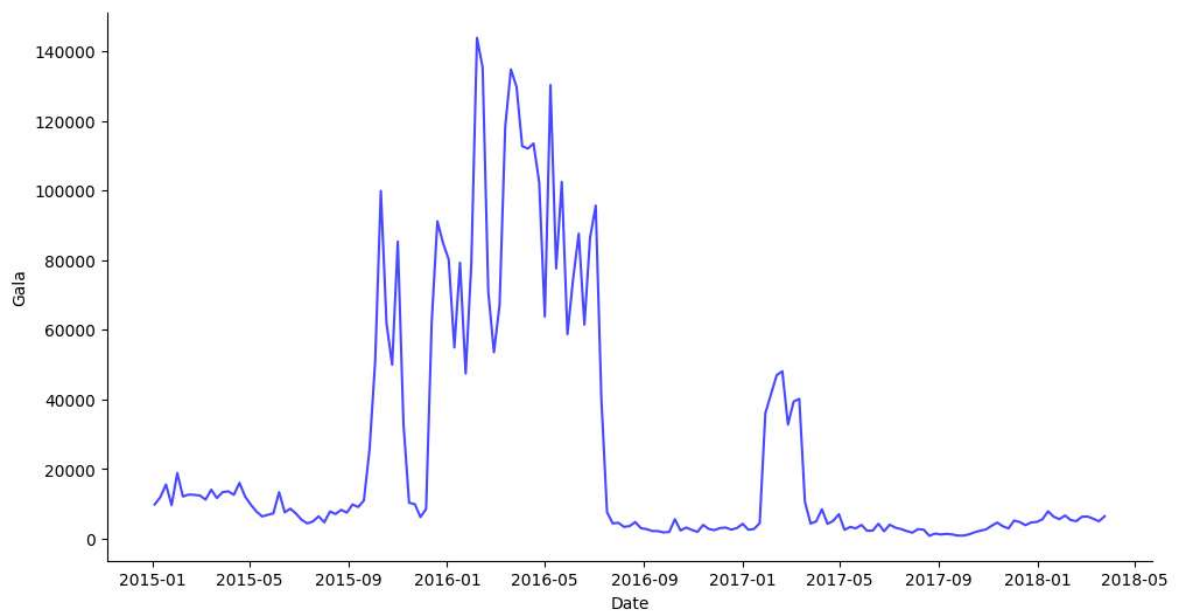
1. เลือกคอลัมน์ Date และ คอลัมน์สายพันธุ์ที่จะนำมาใช้
2. Interpolate เพื่อเติมข้อมูลที่ขาดหายไป (NaN)
3. เปลี่ยน format ของ Date จาก string เป็น datetime64
4. เรียง Date จากเก่าไปหาใหม่

สามารถแสดงผลตัวอย่างได้ดังนี้

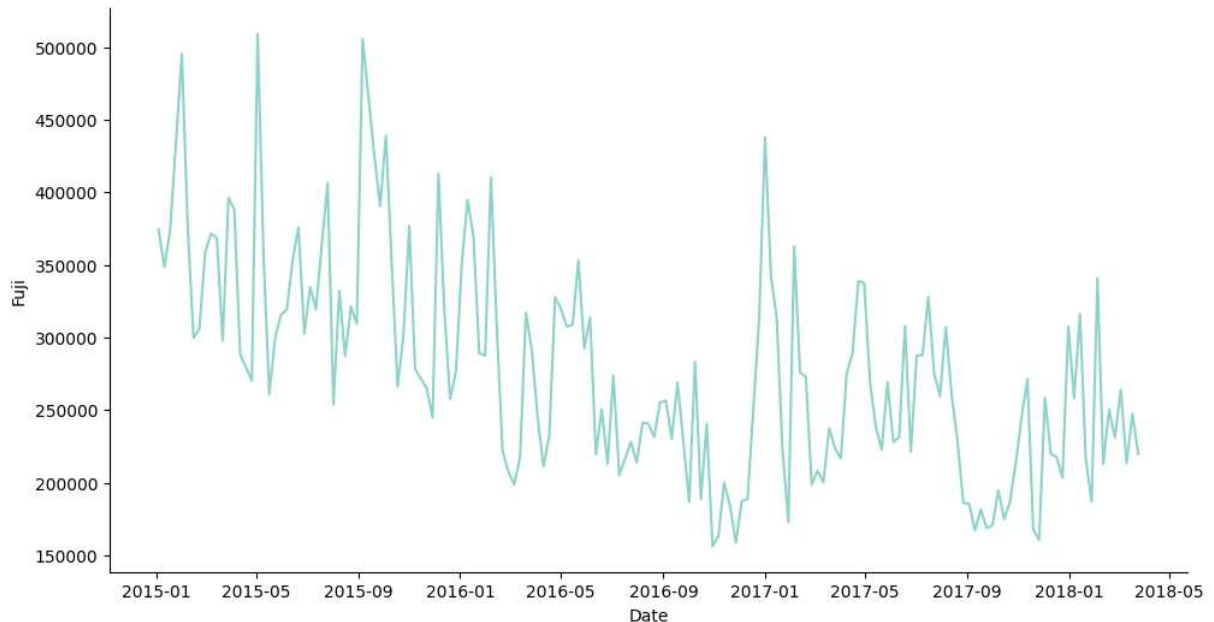
- กราฟของตัวอย่าง Apple สายพันธุ์ Envi



- กราฟของตัวอย่าง Apple สายพันธุ์ Gala



- กราฟของตัวอย่าง Apple สายพันธุ์ Fuji



### เทคนิคที่ใช้

ในการทดลองนี้ได้ทำการทดลองออกเป็น 2 แบบ คือ

1. คาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา
2. คาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา

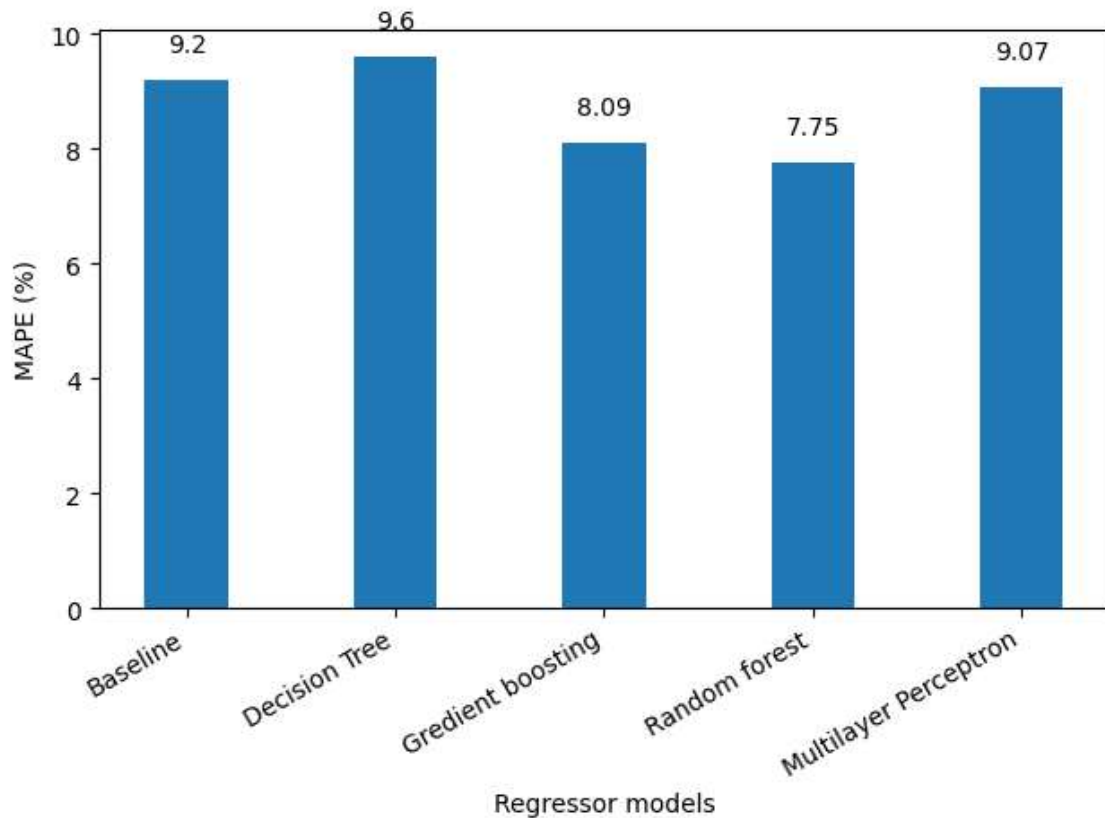
และในการทดลองนี้ได้ใช้โมเดล 4 โมเดล ดังนี้

1. Decision tree regressor
2. Gradient boosting (สำหรับการทดลองแบบที่ 2 จะใช้ RegressorChain ร่วมด้วย)
3. Random Forest มี parameter max\_depth=5
4. Multi-layer Perceptron Regressor มี parameter max\_iter=1000, solver='adam', early\_stopping=True, learning\_rate\_init=0.01

## ผลการทดลอง

### *Apple สายพันธุ์ Envi*

การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา



#### 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 9.60%

#### 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 8.09%

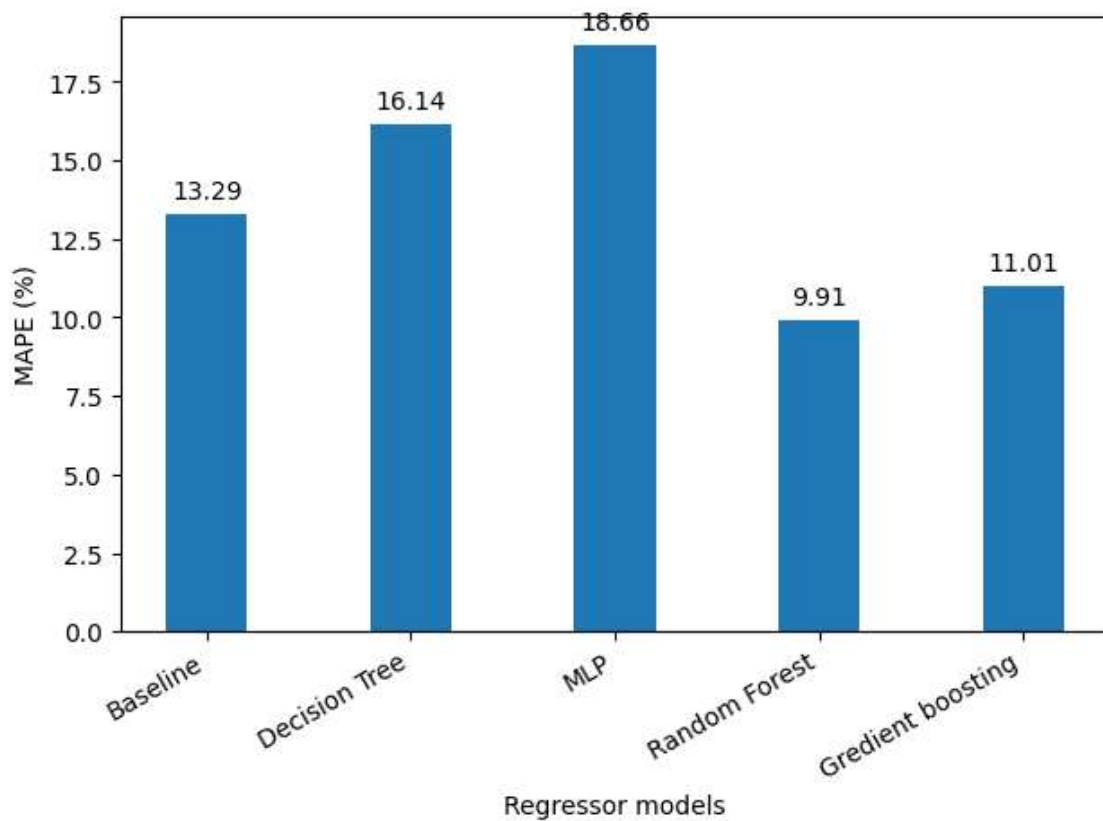
#### 3. Random Forest

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 7.75%

#### 4. Multi-layer Perceptron Regressor

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 9.07%

การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา



##### 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 16.14%

##### 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 11.01%

##### 3. Random Forest

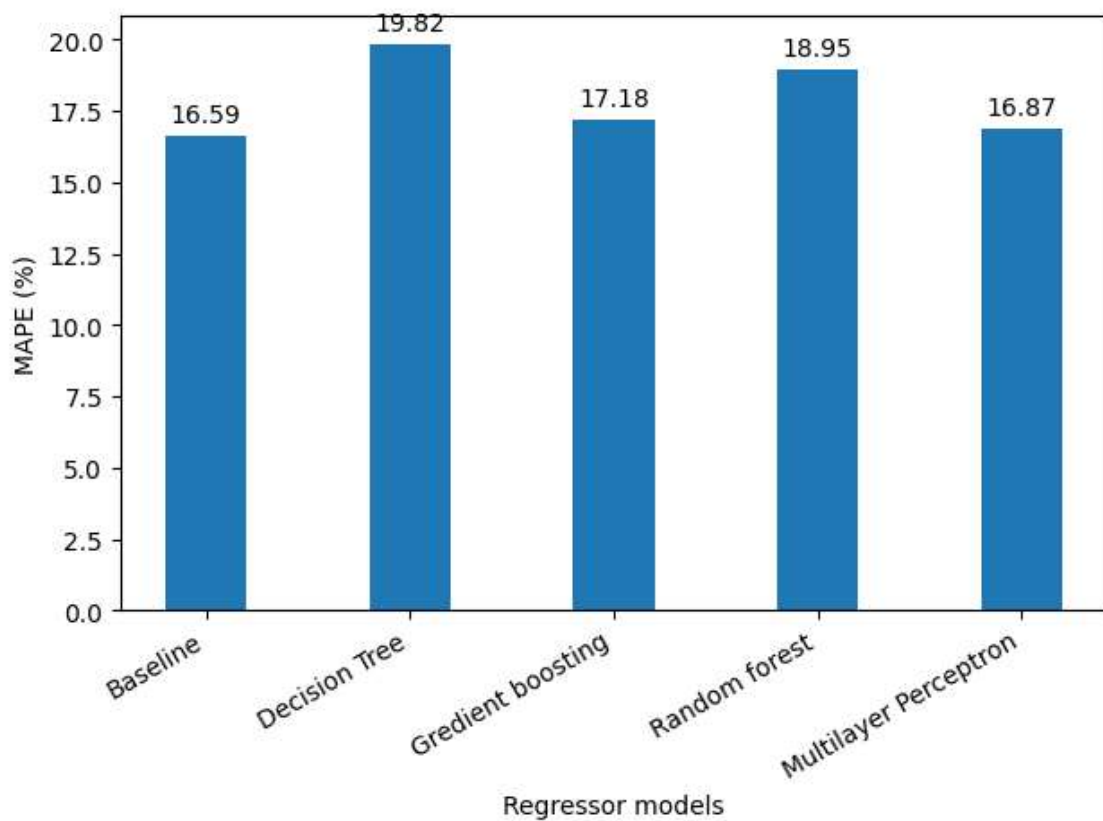
มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 9.91%

#### 4. Multi-layer Perceptron Regressor (MLP)

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 18.66%

#### Apple สายพันธุ์ Fuji

การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา



##### 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 19.82%

##### 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 17.18%

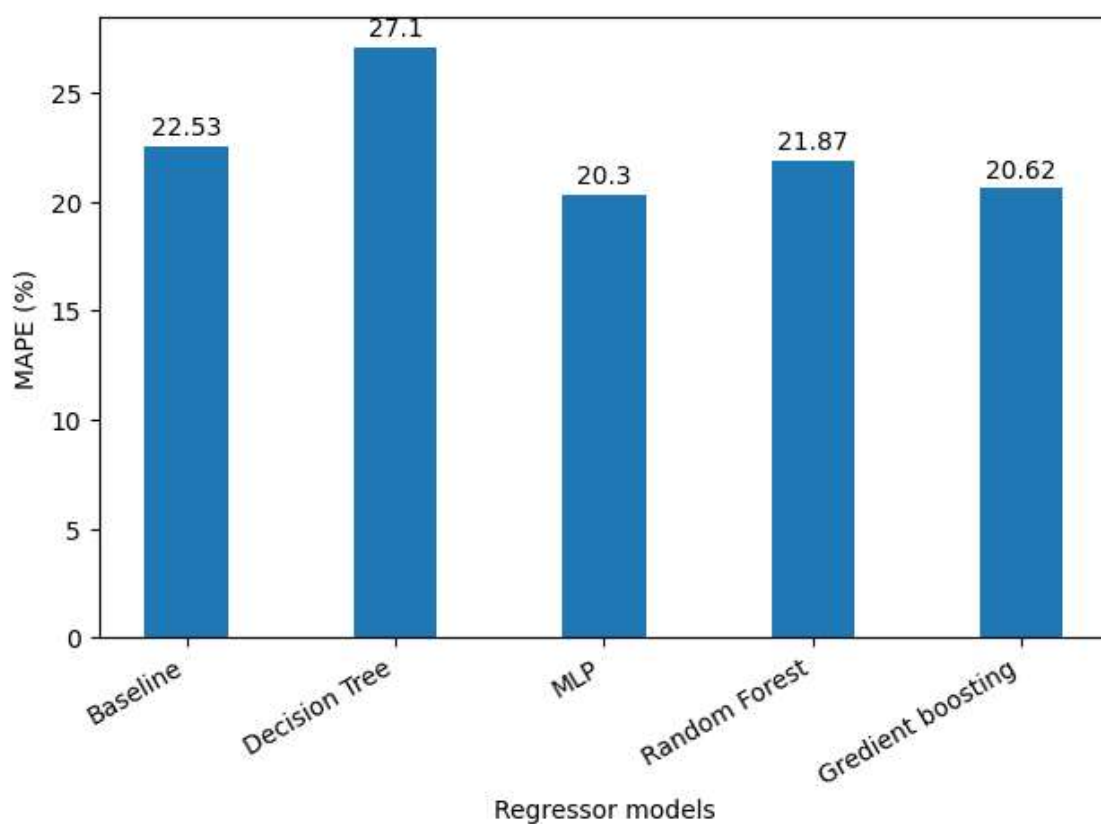
### 3. Random Forest

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 18.95%

### 4. Multi-layer Perceptron Regressor

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 16.87%

การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา



#### 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 27.10%

#### 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 20.62%

### 3. Random Forest

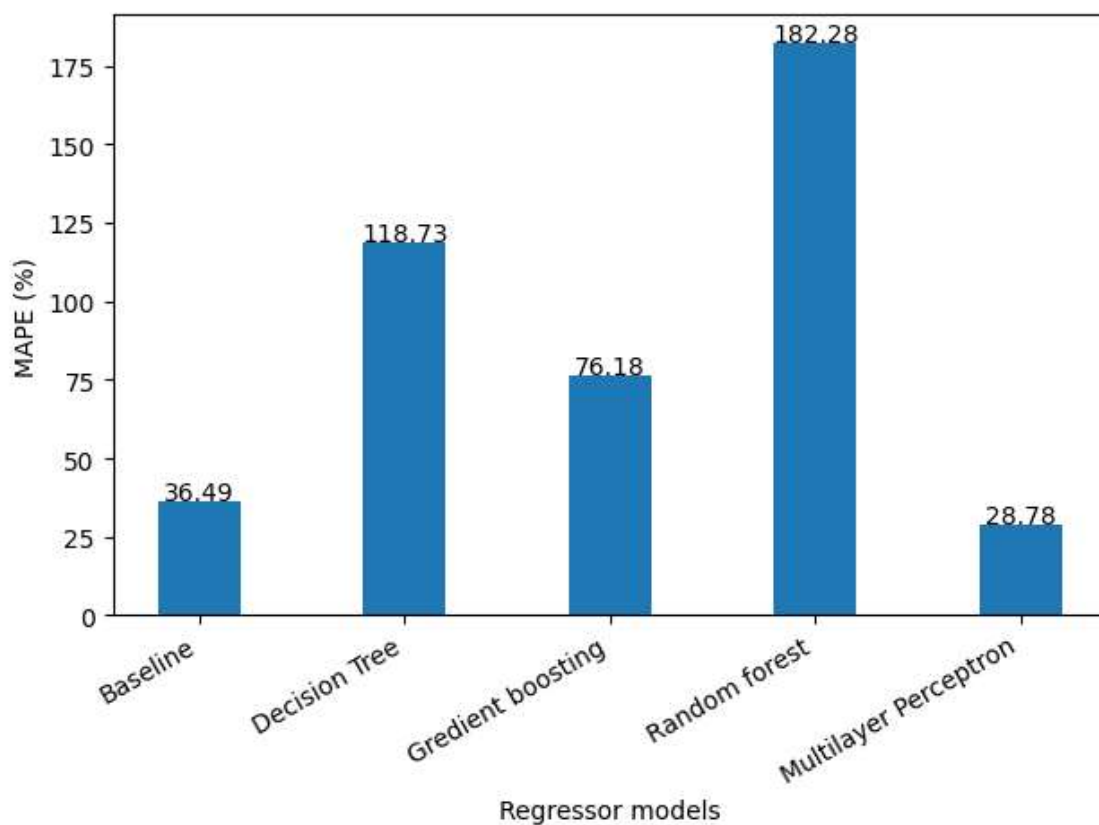
มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 21.87%

### 4. Multi-layer Perceptron Regressor (MLP)

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 20.30%

## Apple สายพันธุ์ Gala

การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา



### 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 118.73%



## 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 76.18%

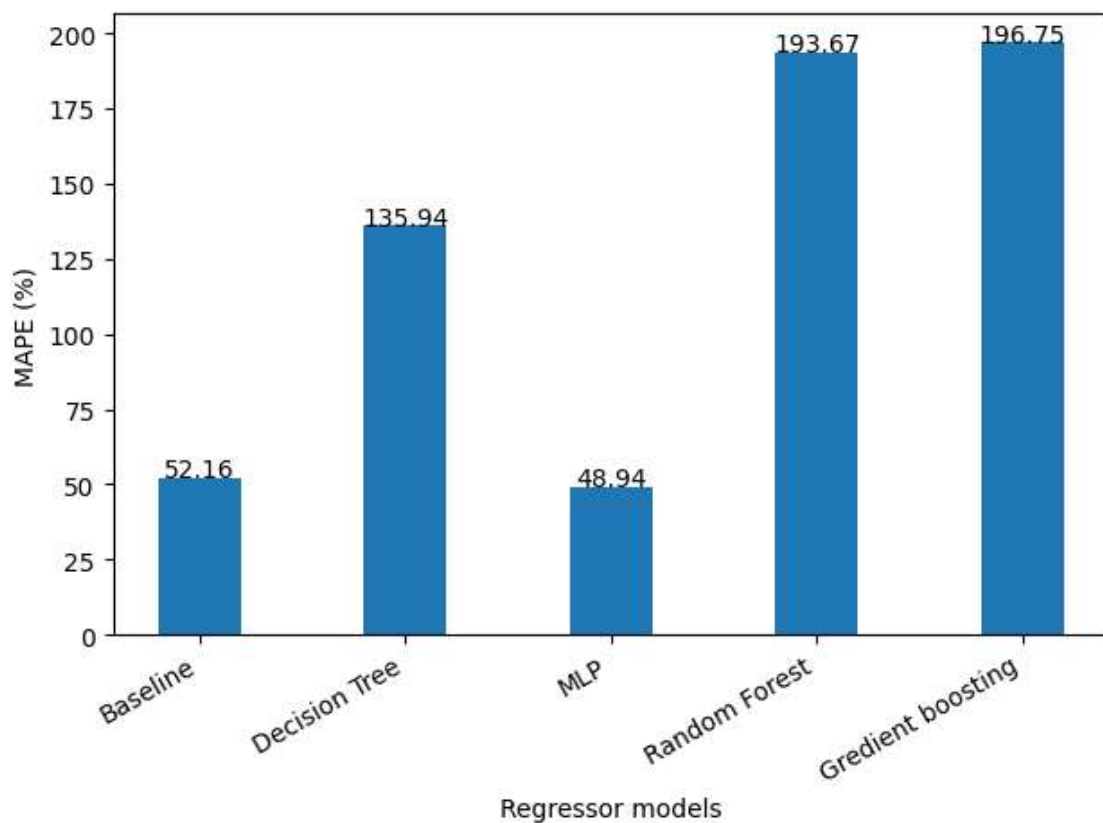
## 3. Random Forest

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 182.28%

## 4. Multi-layer Perceptron Regressor

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 28.78%

การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา



## 1. Decision Tree

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 135.94%

## 2. Gradient Boosting

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 196.75%

## 3. Random Forest

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 193.67%

## 4. Multi-layer Perceptron Regressor (MLP)

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) อยู่ที่ 48.94%

### เปรียบเทียบผลการทดลอง

#### Apple สายพันธุ์ Envi

การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Random Forest (7.75%), Gradient Boosting (8.09%), Multi-layer Perceptron Regressor (9.07%), Decision Tree (9.60%)

การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Random Forest (9.91%), Gradient Boosting (11.01%), Decision Tree (16.14%), Multi-layer Perceptron Regressor (18.66%)

## Apple สายพันธุ์ Fuji

### การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Multi-layer Perceptron Regressor (16.87%), Gradient Boosting (17.18%), Random Forest (18.95%), Decision Tree (19.82%)

### การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Multi-layer Perceptron Regressor (20.30%), Gradient Boosting (20.62%), Random Forest (21.87%), Decision Tree (27.10%)

## Apple สายพันธุ์ Gala

### การคาดการณ์ลำดับถัดไปโดยใช้ลำดับการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Multi-layer Perceptron Regressor (28.78%), Gradient Boosting (76.18%), Decision Tree (118.73%), Random Forest (182.28%)

### การคาดการณ์ลำดับในอนาคตโดยใช้ลำดับของการสังเกตที่ผ่านมา

มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์จากทั้ง 4

โมเดลเรียงลำดับจากค่าที่ดีที่สุด(น้อยไปมาก)ได้ตามนี้ Multi-layer Perceptron Regressor (48.94%), Decision Tree (135.94%), Random Forest (193.67%), Gradient Boosting (196.75%)

### สรุปผลการทดลอง

จากการทดลองผลลัพธ์ทั้ง 3 สายพันธุ์ พบว่า Multi-layer Perceptron Regressor จะมีประสิทธิภาพที่ดีที่สุดในการทำนาย ตามด้วย Random Forest, Gradient Boosting, และ Decision Tree ตามลำดับ ทั้งในส่วนของคาดการณ์ลำดับถัดไปและการคาดการณ์ลำดับในอนาคต

### ชื่อผู้จัดทำ

- นายกษมา เหล็กมาด 6410406495 (หัวหน้ากลุ่ม)
- นายภูเบศ ปานทอง 6410406835
- นายรัตนันริศ จำนง 6410406851