

MODEL
EVALUATION
FOR CLASSIFICATION

BY TAUTOLOGY

Model Evaluation for Classification



Confusion Matrix

Confusion Matrix

เป็น covid \Rightarrow positive
ไม่เป็น covid \Rightarrow negative

Confusion Matrix คือ ตารางแสดงผลลัพธ์ของการ prediction สำหรับปัญหาแบบ classification

| | | Actual Value | |
|---------------|----------|----------------|----------------|
| | | Positive | Negative |
| Predict Value | Positive | True Positive | False Positive |
| | Negative | False Negative | True Negative |

Confusion Matrix

True Negative



ผลลัพธ์ของการพยากรณ์

- True = พยากรณ์ถูก
- False = พยากรณ์ผิด

ค่าที่พยากรณ์ออกมา




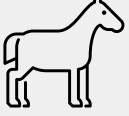

- Positive (ค่าที่เราพิจารณา)
- Negative (ค่าที่เราไม่ได้พิจารณา)

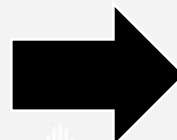
Confusion Matrix





| | | Actual Value | |
|-----------------|----------|------------------------|------------------------|
| | | Positive | Negative |
| Predicted Value | Positive | True Positive (TP) | False Positive (FP) |
| | Negative | False Negative (FN) | True Negative (TN) |

Confusion Matrix

สุนัข => positive
ม้า => negative

| Actual | Predicted |
|---|-----------|
|  | Horse |
|  | Horse |
|  | Dog |
|  | Dog |
|  | Dog |









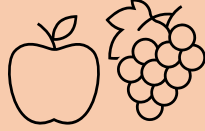

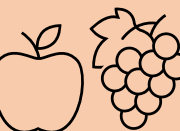
| | | Actual | |
|---------|--|---|---|
| | |  |  |
| Predict |  | 2 TP | 1 FP |
| |  | 1 FN | 1 TN |






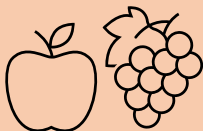
Confusion Matrix




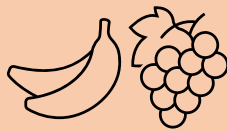
positive (บวก) \Rightarrow Banana
negative \Rightarrow Apple, Grape

| | Actual | Predicted |
|----------|---|-------------------------|
| |  | Banana |
| Negative |  | Apple |
| Negative |  | Apple |
| Negative |  | Grape |
| |  | Grape Banana |

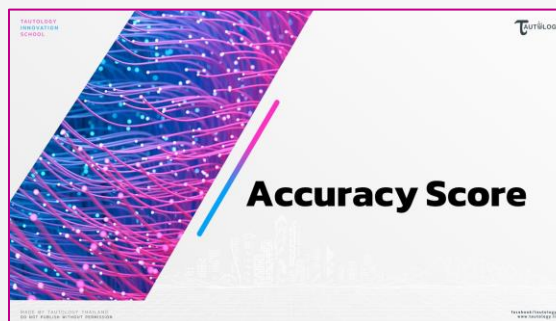
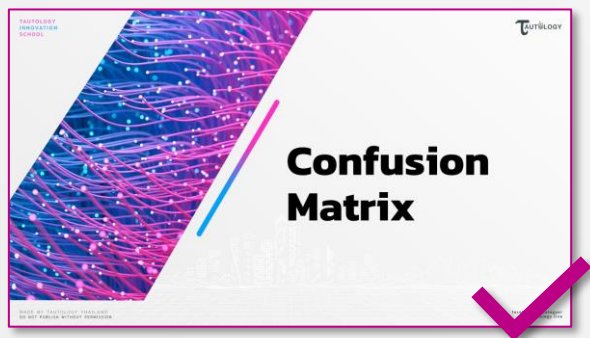
| | | Actual | |
|---------|--|---|---|
| | |  |  |
| Predict |  | 1 TP | 1 FP |
| |  | 0 FN | 3 TN |

Confusion Matrix

| | | Actual | |
|---------|---|---|--|
| | |  |  |
| Predict |  | TP | FP |
| |  | FN | TN |

| | | Actual | |
|---------|---|---|---|
| | |  |  |
| Predict |  | TP | FP |
| |  | FN | TN |

Model Evaluation for Classification








Accuracy Score

Accuracy Score



- What is Accuracy Score?
- Formula
- Step to Calculate Accuracy Score
- Example
- Code
- ข้อควรระวัง



What is Accuracy Score?

Accuracy Score คือ สัดส่วนของข้อมูลที่พยากรณ์ได้ถูกต้อง ต่อข้อมูลทั้งหมด

| Actual | Predicted |
|---|-----------|
|  | Apple |
|  | Banana |
|  | Apple |
|  | Banana |
|  | Apple |



| Actual | |
|---|---|
|  |  |
| 2 ^{TP} | 1 ^{FP} |
| 1 ^{FN} | 1 ^{TN} |

| Predict |  |  |
|---------|--|---|
|---------|--|---|

What is Accuracy Score?

| | | Actual Value | |
|---------------|----------|--------------|----------|
| | | Positive | Negative |
| Predict Value | Positive | TP | FP |
| | Negative | FN | TN |

$$accuracy = \frac{\text{จำนวนที่พยากรณ์ถูก}}{\text{จำนวนข้อมูลทั้งหมด}}$$





Formula

$$accuracy = \frac{TP + TN}{n}$$

- โดยที่
- ◆ TP คือ ค่า True Positive
 - ◆ TN คือ ค่า True Negative
 - ◆ n คือจำนวน sample ทั้งหมด

Formula

ตัวอย่าง

| | | Actual | |
|---------|--|---|--|
| | |  |  |
| Predict |  | 2 ^{TP} | 1 ^{FP} |
| |  | 1 ^{FN} | 1 ^{TN} |

$$accuracy = \frac{\text{จำนวนที่พยากรณ์ถูก}}{\text{จำนวนข้อมูลทั้งหมด}}$$

$$= \frac{TP + TN}{\text{จำนวนข้อมูลทั้งหมด}}$$

$$= \frac{2 + 1}{2 + 0 + 1 + 2} = \frac{3}{5}$$

$$= 0.6$$

Step to Calculate Accuracy Score

1. เก็บค่า y_i และ \hat{y}_i
2. หาค่า TP และ TN ของ model
3. วัดประสิทธิภาพของ model ตามสูตรของ *accuracy*

Example

1. เก็บค่า y_i และ \hat{y}_i

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol

เป็นโรคอ้วน = 1 ไม่เป็นโรค = 0

Example

2. คำนวณค่า TP และ TN

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |



| | | Actual | |
|---------|----------------|-----------------|-----------------|
| | | 1 (เป็นโรค) | 0 (ไม่เป็นโรค) |
| Predict | 1 (เป็นโรค) | 1 ^{TP} | 1 ^{FP} |
| | 0 (ไม่เป็นโรค) | 2 ^{FN} | 3 ^{TN} |

Example

3. วัดประสิทธิภาพของ model ตามสูตรของ *accuracy*

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

$$\begin{aligned} accuracy &= \frac{TP + TN}{n} \\ &= \frac{1 + 3}{7} \\ &= \frac{4}{7} \\ &= 0.5714 \end{aligned}$$

Code

| | Actual_Fat | Predicted_Fat |
|---|------------|---------------|
| 0 | 1.0 | 0.0 |
| 1 | 0.0 | 0.0 |
| 2 | 0.0 | 1.0 |
| 3 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 |
| 5 | 1.0 | 0.0 |
| 6 | 1.0 | 1.0 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol
เป็นโรคอ้วน = 1 ไม่เป็นโรค = 0

Code

```
1 accuracy_score(y_true, y_pred)
```

```
0.5714285714285714
```

ข้อควรระวัง

“เมื่อข้อมูลอยู่ในสถานะ **imbalanced class** จะใช้ **accuracy** อธิบายได้ไม่ดี”

ตัวอย่าง ให้หมอปาลอมตรวจโรคคนไข้จำนวน 100 คน หมอปาลอมตรวจพบว่าคนไข้ไม่เป็นโรคเกือบทั้งหมด

| | | Actual | |
|---------|------------|---------|------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | 1 TP | 0 FP |
| | ไม่เป็นโรค | 7 FN | 92 TN |

๑๑

ข้อควรระวัง

แต่ค่า accuracy ที่คำนวณออกมาได้สูงถึง 0.93

| | | Actual | |
|---------|------------|-------------|--------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | 1 TP | 0 FP |
| | ไม่เป็นโรค | 7 FN | 92 TN |

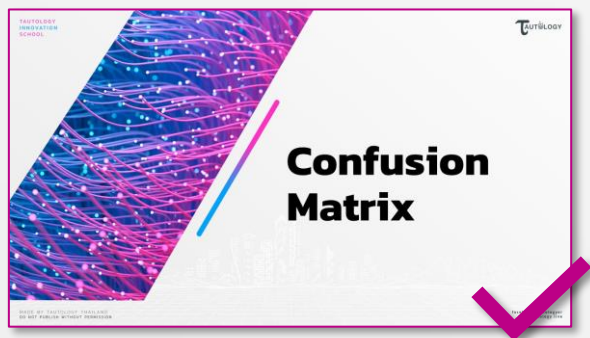
$$accuracy = \frac{TP + TN}{n}$$

$$= \frac{92 + 1}{100}$$

$$= \frac{93}{100} = 0.93 \approx 93\%$$

นั่นหมายความว่า แม้หมอปาลอมจะไม่สามารถตรวจโรคได้จริง แต่การวัดประสิทธิภาพผ่าน accuracy สูงถึง 0.93

Model Evaluation for Classification








Precision Score

Precision Score





- What is Precision Score?
- Formula
- Step to Calculate Precision Score
- Example
- Code

What is Precision Score?

Precision Score คือ สัดส่วนของข้อมูลที่พยากรณ์ว่าเป็น positive ได้อย่างถูกต้อง
ต่อข้อมูลที่พยากรณ์เป็น positive ทั้งหมด

| Actual | Predicted |
|---|-----------|
|  | Apple |
|  | Banana |
|  | Apple |
|  | Banana |
|  | Apple |



| | | Actual | |
|---------|---|---|---|
| | |  |  |
| Predict |  | 2 ^{TP} | 1 ^{FP} |
| |  | 1 ^{FN} | 1 ^{TN} |

What is Precision Score?

| | | Actual Value | |
|---------------|----------|--------------|----------|
| | | Positive | Negative |
| Predict Value | Positive | TP | FP |
| | Negative | FN | TN |

precision

$$= \frac{\text{จำนวนที่พยากรณ์ถูกว่าเป็น } positive}{\text{จำนวนที่พยากรณ์ว่าเป็น } positive \text{ ทั้งหมด}}$$





Formula

$$precision = \frac{TP}{TP + FP}$$

โดยที่ ♦ TP คือ ค่า True Positive
♦ FP คือ ค่า False Positive

Formula

ตัวอย่าง

| | | Actual | |
|---------|--|---|---|
| | |  |  |
| Predict |  | 2 TP | 1 FP |
| |  | 1 FN | 1 TN |

$$\text{precision} = \frac{\text{จำนวนที่พยากรณ์ถูกว่าเป็น apple}}{\text{จำนวนที่พยากรณ์ว่าเป็น apple ทั้งหมด}}$$

$$= \frac{TP}{TP + FP}$$

$$= \frac{2}{2 + 1}$$

$$= 0.67$$

Step to Calculate Precision Score

1. เก็บค่า y_i และ \hat{y}_i
2. หาค่า TP และ FP ของ model
3. วัดประสิทธิภาพของ model ตามสูตรของ *precision*

Example

1. เก็บค่า y_i และ \hat{y}_i

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol

เป็นโรคอ้วน = 1 ไม่เป็นโรค = 0

Example

2. คำนวณค่า TP และ FP

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |



| | | Actual | |
|---------|----------------|-----------------|-----------------|
| | | 1 (เป็นโรค) | 0 (ไม่เป็นโรค) |
| Predict | 1 (เป็นโรค) | 1 ^{TP} | 1 ^{FP} |
| | 0 (ไม่เป็นโรค) | 2 ^{FN} | 3 ^{TN} |

Example

3. วัดประสิทธิภาพของ model ตามสูตรของ *precision*

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP} \\ &= \frac{1}{1 + 1} \\ &= 0.5 \end{aligned}$$

Code

| | Actual_Fat | Predicted_Fat |
|---|------------|---------------|
| 0 | 1.0 | 0.0 |
| 1 | 0.0 | 0.0 |
| 2 | 0.0 | 1.0 |
| 3 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 |
| 5 | 1.0 | 0.0 |
| 6 | 1.0 | 1.0 |

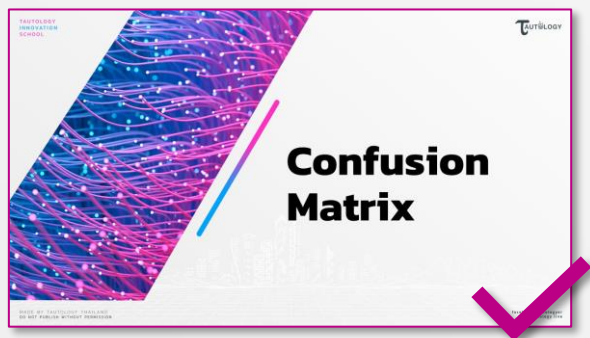
ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol
**เป็นโรคอ้วน = 0 ไม่เป็นโรค = 1

Code

```
1 precision_score(y_true, y_pred)
```

```
0.5
```

Model Evaluation for Classification








Recall Score

Recall Score





- What is Recall Score?
- Formula
- Step to Calculate Recall Score
- Example
- Code
- Precision vs Recall

What is Recall Score?

Recall Score คือ สัดส่วนของข้อมูลที่พยากรณ์ว่าเป็น positive ได้อย่างถูกต้อง ต่อ ข้อมูลที่ actual เป็น positive ทั้งหมด

| Actual | Predicted |
|---|-----------|
|  | Apple |
|  | Banana |
|  | Apple |
|  | Banana |
|  | Apple |



| | | Actual | |
|---------|---|--|--|
| Predict |  |  2 TP |  1 FP |
| |  | 1 FN | 1 TN |

What is Recall Score?

| | | Actual Value | |
|---------------|----------|--------------|----------|
| | | Positive | Negative |
| Predict Value | Positive | TP | FP |
| | Negative | FN | TN |

Recall

$$= \frac{\text{จำนวนที่พยากรณ์ถูกว่าเป็น } positive}{\text{จำนวนที่ค่าจริงเป็น } positive \text{ ทั้งหมด}}$$

Formula





$$recall = \frac{TP}{TP + FN}$$

↳ actual ทั้งหมดที่เป็น positive

- โดยที่
- ◆ TP คือ ค่า True Positive
 - ◆ FN คือ ค่า False Negative

Formula

ตัวอย่าง

| | | Actual | |
|---------|--|--|--|
| Predict |  |  2 TP |  1 FP |
| |  | 1 FN | 1 TN |

$$Recall = \frac{\text{จำนวนที่พยากรณ์ถูกว่าเป็น } apple}{\text{จำนวนที่ค่าจริงเป็น } apple \text{ ทั้งหมด}}$$

$$= \frac{TP}{TP + FN}$$

$$= \frac{2}{2 + 1}$$

$$= 0.67$$

Step to Calculate Recall Score

1. เก็บค่า y_i และ \hat{y}_i
2. หาค่า TP และ FN
3. วัดประสิทธิภาพของ model ตามสูตรของ *recall*

Example

1. เก็บค่า y_i และ \hat{y}_i

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol

เป็นโรคอ้วน = 1 ไม่เป็นโรค = 0

Example

2. คำนวณค่า TP และ FN

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |



| | | Actual | |
|---------|----------------|---------------------------------------|--|
| Predict | 1 (เป็นโรค) | 1 (เป็นโรค) 1 ^{TP} | 0 (ไม่เป็นโรค) 1 ^{FP} |
| | 0 (ไม่เป็นโรค) | 2 ^{FN} | 3 ^{TN} |

Example

3. วัดประสิทธิภาพของ model ตามสูตรของ *recall*

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

$$\begin{aligned} recall &= \frac{TP}{TP + FN} \\ &= \frac{1}{1 + 2} \\ &= 0.33 \end{aligned}$$

Code

| | Actual_Fat | Predicted_Fat |
|---|------------|---------------|
| 0 | 1.0 | 0.0 |
| 1 | 0.0 | 0.0 |
| 2 | 0.0 | 1.0 |
| 3 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 |
| 5 | 1.0 | 0.0 |
| 6 | 1.0 | 1.0 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol
**เป็นโรคอ้วน = 0 ไม่เป็นโรค = 1

Code

```
1 recall_score(y_true, y_pred)
```

```
0.3333333333333333
```


Precision vs Recall

Precision VS Recall

>> จากการพยากรณ์ว่าเป็น positive ทั้งหมด มีจำนวนที่พยากรณ์ถูกเท่าไร

$$>> precision = \frac{TP}{TP+FP}$$

>> จากข้อมูลจริงที่เป็น positive ทั้งหมด มีจำนวนที่พยากรณ์ถูกเท่าไร

$$>> recall = \frac{TP}{TP+FN}$$

Precision vs Recall

**Precision และ Recall
ควรใช้เมื่อไหร่ ?**

Precision vs Recall

▪ Precision

ตัวอย่าง ในการพิจารณาคดี จำเป็นที่จะต้องคำนึงว่า ผู้ที่ไม่ได้กระทำความผิด ไม่ควรได้รับโทษ

กำหนดให้

- การกระทำความผิด => positive
- ไม่ได้กระทำความผิด => negative

Precision vs Recall

■ Precision

| | | Actual | |
|---------|-------------|--------|-------------|
| | | ทำผิด | ไม่ได้ทำผิด |
| Predict | ทำผิด | TP | FP |
| | ไม่ได้ทำผิด | FN | TN |

$$precision = \frac{TP}{TP + FP}$$

“ในกรณีนี้ เราจำเป็นต้องพิจารณา **precision** ให้มีค่าสูง ๆ เพราะ False Positive (คนที่ไม่ได้ทำผิดแต่ถูกพยากรณ์ว่าทำผิด) จะมีค่าน้อย ๆ”

Precision vs Recall

▪ Recall

ตัวอย่าง ในการตรวจโรคมะเร็งเพื่อเข้ารับการรักษ่า จำเป็นที่จะต้องคำนึงว่า ผู้ป่วยเป็น มะเร็งทุกคนนั้นจะต้องได้รับการรักษ่า กำหนดให้

- เป็นโรคมะเร็ง => positive
- ไม่เป็นโรคมะเร็ง => negative

Precision vs Recall

▪ Recall

| | | Actual | |
|---------|------------|---------|------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | TP | FP |
| | ไม่เป็นโรค | FN | TN |

$$recall = \frac{TP}{TP + FN}$$

“ในกรณีนี้ เราจำเป็นต้องพิจารณา **recall** ให้มีค่าสูง ๆ เพราะ False Negative (คนที่เป็นโรคแต่ถูกพยากรณ์ว่าไม่เป็นโรค) จะมีค่าน้อย ๆ”

Model Evaluation for Classification



F1 Score

F1 Score

- What is F1 Score?
- Formula
- Step to Calculate F1 Score
- Example
- Code
- F1 Score with Imbalanced Class

What is F1 Score?

F1 Score คือ ค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall

Precision

Recall



What is F1 Score?

Harmonic mean เป็นการหาค่าเฉลี่ยประเภทหนึ่ง โดยการหาค่าเฉลี่ยวิธีนี้จะ

- ให้น้ำหนักน้อยกับข้อมูลที่มีค่าเยอะ
- ให้น้ำหนักเยอะกับข้อมูลที่มีค่าน้อย

What is F1 Score?

ให้น้ำหนัก**เยอะ**กับข้อมูลที่มีค่าน้อย

mean of: [37, 35, 40, 35, 29, 51, 31, 33, 34, 30, 29, 33, 37, 36, 0.01]

◆ Harmonic Mean = 0.14939025281869237

◆ Arithmetic Mean = 32.66733333333333

What is F1 Score?

Harmonic mean สามารถเขียนให้อยู่ในรูปดังต่อไปนี้

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

โดยที่

- ♦ n คือจำนวนข้อมูลทั้งหมด
- ♦ x_i คือข้อมูลที่ i

What is F1 Score?

ตัวอย่าง กำหนดให้ $x = \{10, 100\}$

Harmonic mean จะสามารถคำนวณได้ดังต่อไปนี้

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{2}{\frac{1}{10} + \frac{1}{100}} = \frac{2}{0.11} = 18.18$$

What is F1 Score?

F1 Score จะพิจารณาค่าระหว่าง precision และ recall ผ่าน harmonic mean ซึ่งสามารถจัดรูปได้ดังนี้

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

maximum $\Rightarrow 1$

minimum \Rightarrow

What is F1 Score?

โดย F1 score จะสามารถตีความได้ดังต่อไปนี้

- ถ้าค่า F1 score **มาก** หมายความว่า ค่า precision และ recall มีค่ามากทั้งคู่
- ถ้าค่า F1 score **น้อย** หมายความว่า ค่า precision และ recall มีค่าน้อยทั้งคู่ หรือมีค่าใดค่าหนึ่งน้อย

Formula

$$F1 = 2 \times \frac{precision * recall}{precision + recall}$$

โดยที่ ♦ $precision = \frac{TP}{TP+FP}$

♦ $recall = \frac{TP}{TP+FN}$

Step to calculate F1 Score

1. เก็บค่า y_i และ \hat{y}_i
2. หาค่า TP, FP และ FN ของ model
3. หาค่า *precision* ของ model
4. หาค่า *recall* ของ model
5. วัดประสิทธิภาพของ model ตามสูตรของ $F1$

Example

1. เก็บค่า y_i และ \hat{y}_i

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol

เป็นโรคอ้วน = 1 ไม่เป็นโรค = 0

Example

2. คำนวณค่า TP , FP และ FN

| | y_i | \hat{y}_i |
|---|-------|-------------|
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 2 | 0 | 1 |
| 3 | 0 | 0 |
| 4 | 0 | 0 |
| 5 | 1 | 0 |
| 6 | 1 | 1 |



| | | Actual | |
|---------|----------------|-----------------|-----------------|
| | | 1 (เป็นโรค) | 0 (ไม่เป็นโรค) |
| Predict | 1 (เป็นโรค) | 1 ^{TP} | 1 ^{FP} |
| | 0 (ไม่เป็นโรค) | 2 ^{FN} | 3 ^{TN} |

Example

3. หาค่า *precision*

$$precision = \frac{TP}{TP + FP}$$

$$= \frac{1}{1 + 1}$$

$$= \frac{1}{2}$$

Example

4. ค่า *recall*

$$recall = \frac{TP}{TP + FN}$$

$$= \frac{1}{1 + 2}$$

$$= \frac{1}{3}$$

Example

5. วัดประสิทธิภาพของ model ตามสูตรของ $F1$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$= 2 \times \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} + \frac{1}{3}}$$

$$= \frac{2}{5}$$

$$= 0.4$$

Code

| | Actual_Fat | Predicted_Fat |
|---|------------|---------------|
| 0 | 1.0 | 0.0 |
| 1 | 0.0 | 0.0 |
| 2 | 0.0 | 1.0 |
| 3 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 |
| 5 | 1.0 | 0.0 |
| 6 | 1.0 | 1.0 |

ตารางแสดงข้อมูลของคนที่เป็นโรคอ้วนจากข้อมูลจริง และการพยากรณ์ที่ได้จากโมเดล
โดยใช้ feature ที่ใช้คือค่า BMI และ Cholesterol
**เป็นโรคอ้วน = 0 ไม่เป็นโรค = 1

Code

```
1 f1_score(y_true, y_pred)
```

```
0.4
```

F1 Score with Imbalanced Class

F1 Score สามารถใช้วัดประสิทธิภาพของ model ที่เป็น imbalanced class ได้ดี

ตัวอย่าง ให้หมอปาลอมตรวจโรคคนไข้จำนวน 100 คน หมอปาลอมตรวจพบว่าคนไข้ไม่เป็นโรคเกือบทั้งหมด

| | | Actual | |
|---------|------------|---------|------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | 1 TP | 0 FP |
| | ไม่เป็นโรค | 7 FN | 92 TN |

F1 Score with Imbalanced Class

- คำนวณผ่าน **accuracy** จะได้ค่าสูงถึง **0.93**

| | | Actual | |
|---------|------------|-------------|--------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | 1 TP | 0 FP |
| | ไม่เป็นโรค | 7 FN | 92 TN |

$$\begin{aligned} accuracy &= \frac{TP + TN}{n} \\ &= \frac{92 + 1}{100} \\ &= \frac{93}{100} = 0.93 \end{aligned}$$

F1 Score with Imbalanced Class

- คำนวณผ่าน **F1 score** จะได้ค่าออกมาเพียง **0.22**

| | | Actual | |
|---------|------------|---------|------------|
| | | เป็นโรค | ไม่เป็นโรค |
| Predict | เป็นโรค | 1 TP | 0 FP |
| | ไม่เป็นโรค | 7 FN | 92 TN |

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$= 2 \times \frac{1 \times \frac{1}{8}}{1 + \frac{1}{8}}$$

$$= 0.22$$

Model Evaluation for Classification



Conclusion

Conclusion

| Name | Formula | How to use |
|-----------|--|--|
| Accuracy | $accuracy = \frac{TP + TN}{n}$ | <ul style="list-style-type: none">ใช้กับการวัด model ที่ไม่เป็น imbalanced class (balanced class) |
| Precision | $precision = \frac{TP}{TP + FP}$ | <ul style="list-style-type: none">ใช้วัด model ที่ต้องการให้เกิด False Positive น้อยๆ |
| Recall | $recall = \frac{TP}{TP + FN}$ | <ul style="list-style-type: none">ใช้วัด model ที่ต้องการให้เกิด False Negative น้อยๆ |
| F1 | $F1 = 2 \times \frac{precision \times recall}{precision + recall}$ | <ul style="list-style-type: none">ใช้วัด model ที่เป็น imbalanced class ได้ต้องการพิจารณาทั้ง precision และ recall พร้อมกัน |

Model Evaluation for Classification

