EVALUATION

BY TAUTOLOGY



Model Evaluation in Classification

- Confusion Matrix
- Accuracy Score
- Precision Score
- Recall Score
- F1 Score
- Conclusion







Confusion Matrix คือ ตารางแสดงผลลัพธ์ของการ prediction สำหรับปัญหาแบบ classification

		Actual Value	
		Positive	Negative
Predict Value	Positive	True Positive	False Positive
Pred Val	Negative	False Negative	True Negative



True Negative





ผลลัพธ์ของการพยากรณ์

- True = พยากรณ์ถูก
- False = พยากรณ์ผิด

ค่าที่พยากรณ์ออกมา

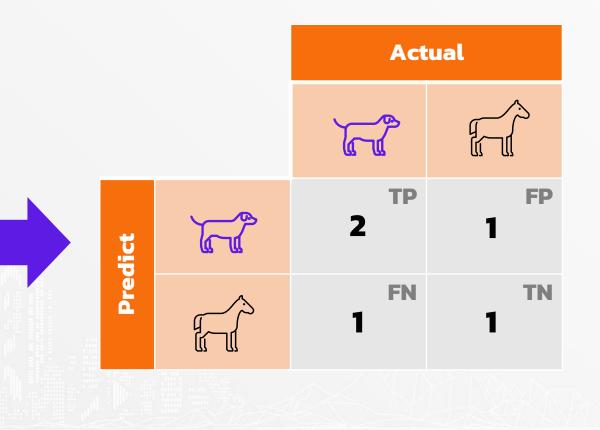
- positive (ค่าที่เราพิจาณา)
- negative (ค่าที่เราไม่ได้พิจาณา)



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

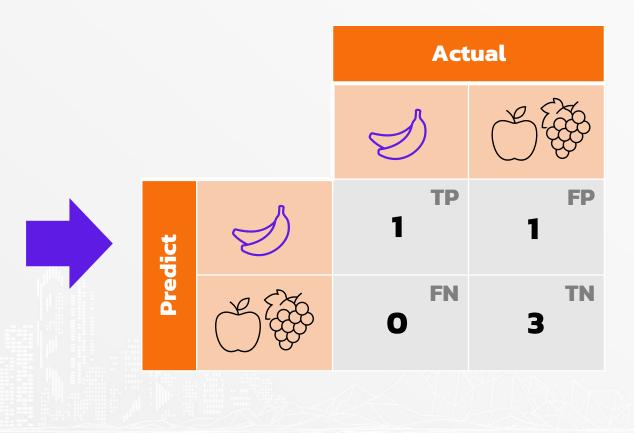


Actual	Predicted
	Horse
	Horse
	Dog
	Dog
	Dog

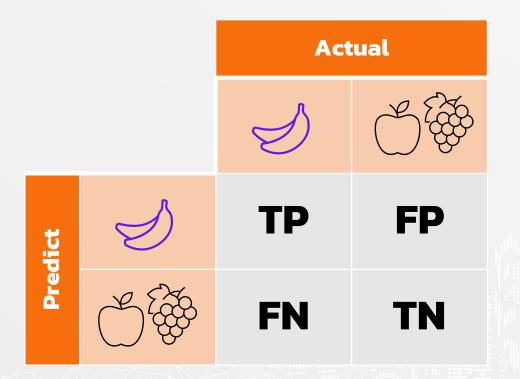


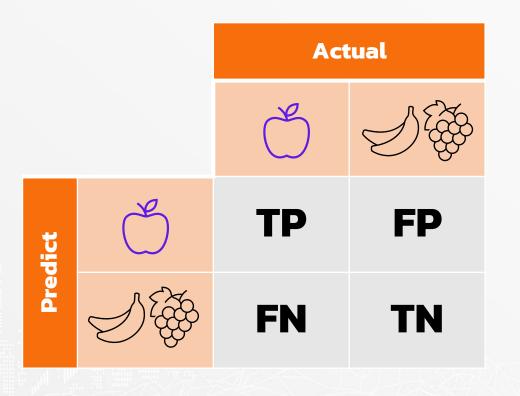


Actual	Predicted
	Banana
7	Apple
Š	Apple
- 	Grape
	Grape











Model Evaluation in Classification

- **✓• Confusion Matrix**
 - Accuracy Score
 - Precision Score
 - Recall Score
 - F1 Score
 - Conclusion





Accuracy Score

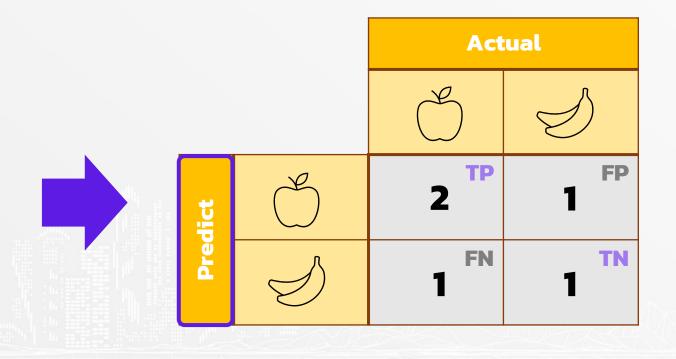
- What is Accuracy Score?
- Formula
- Step to calculate Accuracy Score
- Example
- Code
- ข้อควรระวัง



What is Accuracy Score?

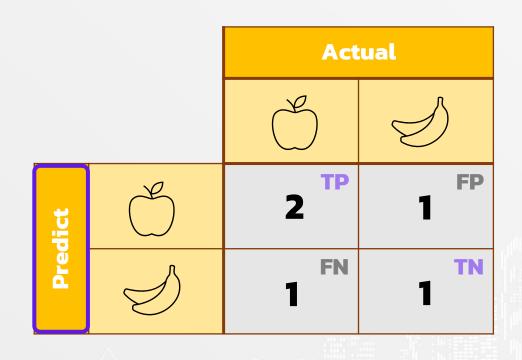
Accuracy Score คือ สัดส่วนของข้อมูลที่พยากรณ์ได้ถูกต้อง ต่อข้อมูลทั้งหมด

Actual	Predicted
	Apple
Š	Banana
Š	Apple
\mathcal{J}	Banana
	Apple





What is Accuracy Score?



$$accuracy = \frac{$$
จำนวนที่พยากรณ์ถูก จำนวนข้อมูลทั้งหมด

$$=rac{TP+TN}{ ext{ จำนวนข้อมูลทั้งหมด}}$$

$$= \frac{2+1}{2+0+1+2}$$
$$= 0.6$$



What is Accuracy Score?

		Actual Value	
		Positive	Negative
Predict Value	Positive	TP	FP
Pred Val	Negative	FN	TN

$$accuracy = \frac{\mathring{\text{-} outunder}}{\mathring{\text{-} outunder}}$$



Formula

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

โดยที่ ◆ TP คือ ค่า True Positive

- ◆ FN คือ ค่า False Negative
- ♦ n คือจำนวน sample ทั้งหมด



Step to calculate Accuracy Score

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



1. เก็บค่า y_i และ \hat{y}_i

	${\bf y_i}$	$\widehat{\boldsymbol{y}}_{i}$
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**



2. คำนวณค่า TP และ TN

	y_i	\widehat{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)	o (ไม่เป็น โรค)
lict	1 (เป็นโรค)	1 TP	1
Predict	o (ไม่เป็น โรค)	2 FN	3 TN



3. วัดประสิทธิภาพของ model ตามสูตรของ accuracy

	y_i	\widehat{y}_i
0	1	0
1	0	0
2	0	1
3	0	0
4	0	0
5	1	0
6	1	1

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

$$= \frac{1+3}{7}$$

$$= \frac{4}{7}$$

$$= 0.5714$$



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 จะไม่สามารถใช้ accuracy อธิบายได้"

ตัวอย่าง ให้หมอปลอมตรวจโรคคนไข้จำนวน 100 คน <u>หมอปลอมตรวจพบว่าคนไข้ไม่</u> เป็นโรคเกือบทั้งหมด

		Actual	
		เป็นโรค	ไม่เป็นโรค
Predict	เป็นโรค	1 TP	OFP
Pre	ไม่เป็นโรค	7 FN	92 ^{TN}



ข้อควรระวัง

แต่ค่า accuracy ที่คำนวณออกมาได้สูงถึง 0.93

		Actual	
		เป็นโรค	ไม่เป็นโรค
Predict	เป็นโรค	1 TP	OFP
	ไม่เป็นโรค	7 FN	92 ^{TN}

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

<mark>นั่นหมายความว่า</mark> แม้หมอปลอมจะไม่สามารถตรวจโรคได้จริง แต่การวัดประสิทธิภาพ ผ่าน accuracy สูงถึง 0.93



Model Evaluation in Classification

- **✓• Confusion Matrix**
- **✓• Accuracy Score**
 - Precision Score
 - Recall Score
 - F1 Score
 - Conclusion





Precision Score

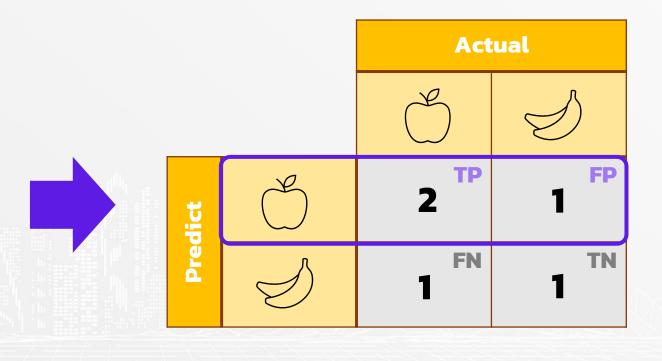
- What is Precision Score?
- Formula
- Step to calculate Precision Score
- Example
- Code



What is Precision Score?

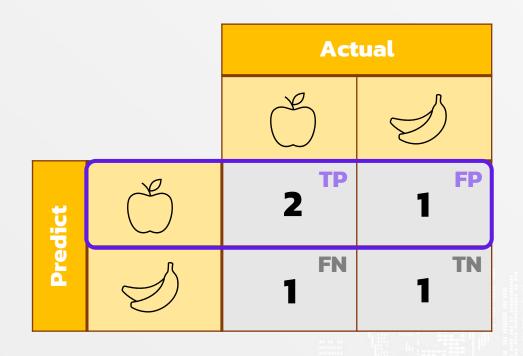
Precision Score คือ สัดส่วนของข้อมูลที่พยากรณ์ได้อย่างถูกต้อง ต่อข้อมูลที่ พยากรณ์เป็น positive ทั้งหมด

Actual	Predicted
\mathcal{J}	Apple
Š	Banana
Š	Apple
	Banana
	Apple





What is Precision Score?



$$precision = rac{ }{ }$$
 จำนวนที่พยากรณ์ถูกว่าเป็น $apple$ ทั้งหมด $= rac{TP}{TP+TN} = rac{2}{2+1} = 0.67$



What is Precision Score?

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

precision

จำนวนที่พยากรณ์ถูกว่าเป็น positive

จำนวนที่พยากรณ์ว่าเป็น postive ทั้งหมด



Formula

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

โดยที่ ◆ TP คือ ค่า True Positive

◆ FN คือ ค่า False Negative



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



1. เก็บค่า y_i และ \hat{y}_i

	y_i	$\widehat{\boldsymbol{y}}_{i}$
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**



2. คำนวณค่า TP และ TN

	y_i	\widehat{y}_i
0	1	0
1	0	0
2	0	1
3	0	О
4	0	0
5	1	0
6	1	1

			Actual	
			1 (เป็นโรค)	o (ไม่เป็น โรค)
	Predict	1 (เป็นโรค)	1 TP	fP 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Pre	o (ไม่เป็น โรค)	2 FN	3 TN



3. วัดประสิทธิภาพของ model ตามสูตรของ precision

	y_i	\widehat{y}_i
0	1	0
1	0	0
2	0	1
3	0	0
4	0	0
5	1	0
6	1	1

$$precision = \frac{TP}{TP + TN}$$
$$= \frac{1}{1+1}$$
$$= 0.5$$



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 in Classification

- **✓** Confusion Matrix
- **✓• Accuracy Score**
- **✓• Precision Score**
 - Recall Score
 - F1 Score
 - Conclusion





Recall Score

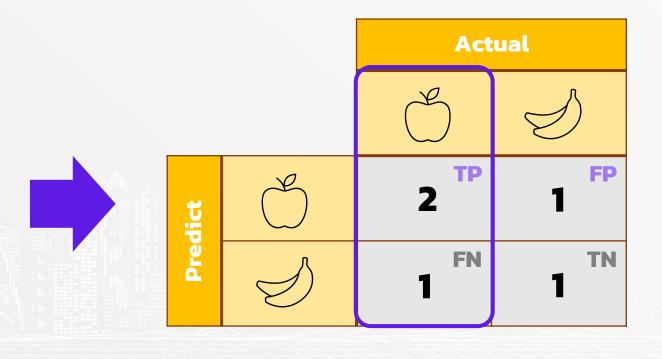
- What is Recall Score?
- Formula
- Step to calculate Recall Score
- Example
- Code
- Precision vs Recall



What is Recall Score?

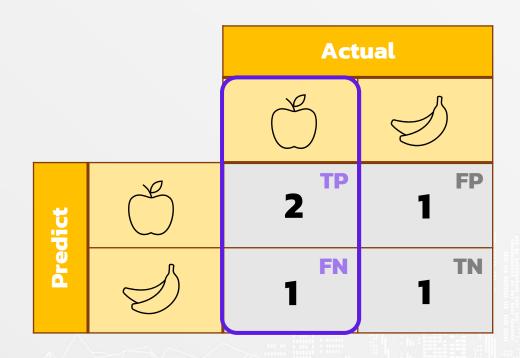
Recall Score คือ สัดส่วนของข้อมูลที่พยากรณ์ได้อย่างถูกต้อง ต่อข้อมูลที่ actual เป็น positive ทั้งหมด

Actual	Predicted
\checkmark	Apple
Š	Banana
	Apple
	Banana
	Apple





What is Recall Score?



$$Recall = rac{$$
จำนวนที่พยากรณ์ถูกว่าเป็น $apple$ $rac{}{}$ จำนวนที่ค่าจริงเป็น $apple$ ทั้งหมด $rac{}{}$

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

$$= \frac{2}{2+1} \\ = 0.67$$



What is Recall Score?

		Actual Value	
		Positive	Negative
Predict Value	Positive	TP	FP
Pred Val	Negative	FN	TN

Recall

จำนวนที่พยากรณ์ถูกว่าเป็น positive

จำนวนที่ค่าจริงเป็น *positive* ทั้งหมด



Formula

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

โดยที่ ◆ TP คือ ค่า True Positive

◆ FN คือ ค่า False Negative



Step to calculate Recall Score

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



1. เก็บค่า y_i และ \hat{y}_i

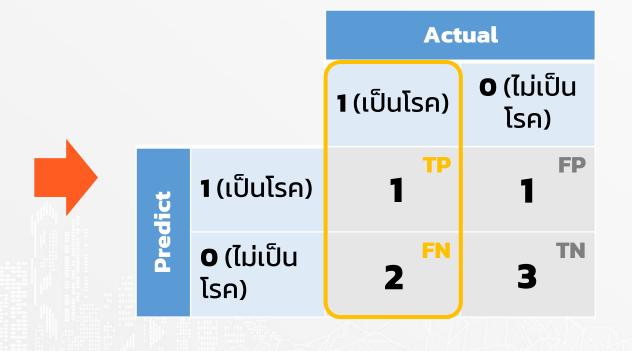
	${\bf y_i}$	$\widehat{\boldsymbol{y}}_{i}$
0	1	0
1	0	0
2	0	1
3	0	0
4	0	0
5	1	0
6	1	1

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



2. คำนวณค่า TP และ FN

	y_i	\widehat{y}_i
0	1	0
1	0	0
2	0	1
3	0	0
4	0	0
5	1	0
6	1	1





3. วัดประสิทธิภาพของ model ตามสูตรของ recall

	y_i	\widehat{y}_i
0	1	0
1	0	0
2	0	1
3	0	0
4	0	0
5	1	0
6	1	1

$$recall = \frac{TP}{TP + FN}$$
$$= \frac{1}{1+2}$$
$$= 0.33$$



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

>> จากการพยากรณ์ว่าเป็น positive ทั้งหมด มีจำนวนที่ พยากรณ์ถูกเท่าไร

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



>> จากข้อมูลจริงที่เป็น positive ทั้งหมด มีจำนวนที่พยากรณ์ถูกเท่าไร

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



Precision และ Recall ควรใช้เมื่อไหร่ ?



Precision

ตัวอย่าง ในการพิจารณาคดี จำเป็นที่จะต้องคำนึงว่า ผู้ที่ไม่ได้กระทำความผิด ไม่ควร ได้รับโทษ

กำหนดให้

- การกระทำความผิด => positive
- ไม่ได้กระทำความผิด => negative



Precision

		Actual	
		ทำผิด	ไม่ได้ทำผิด
dict	ทำผิด	TP	FP
Predict	ไม่ได้ทำผิด	FN	TN

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

"ในกรณีนี้ เราจำเป็นต้องพิจารณา precision ให้มีค่าสูง ๆ เพราะ False Positive (คนที่ไม่ได้ทำผิดแต่ถูกพยากรณ์ว่า ทำผิด) จะมีค่าน้อย ๆ "



Recall

ตัวอย่าง ในการตรวจโรคมะเร็งเพื่อเข้ารับการรักษา จำเป็นที่จะต้องคำนึงว่า ผู้ป่วยเป็น มะเร็งทุกคนนั้นจะต้องได้รับการรักษา

กำหนดให้

- เป็นโรคมะเร็ง => positive
- ไม่เป็นโรคมะเร็ง => negative



Recall

		Actual	
		เป็นโรค	ไม่เป็นโรค
dict	เป็นโรค	TP	FP
Predict	ไม่เป็นโรค	FN	TN

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

"ในกรณีนี้ เราจำเป็นต้องพิจารณา recall ให้มีค่าสูง ๆ เพราะ False Negative (คนที่เป็นโรคแต่ถูกพยากรณ์ว่าไม่ เป็นโรค) จะมีค่าน้อย ๆ"



Model Evaluation in Classification

- **✓** Confusion Matrix
- Accuracy Score
- **✓• Precision Score**
- **✓• Recall Score**
 - F1 Score
 - Conclusion





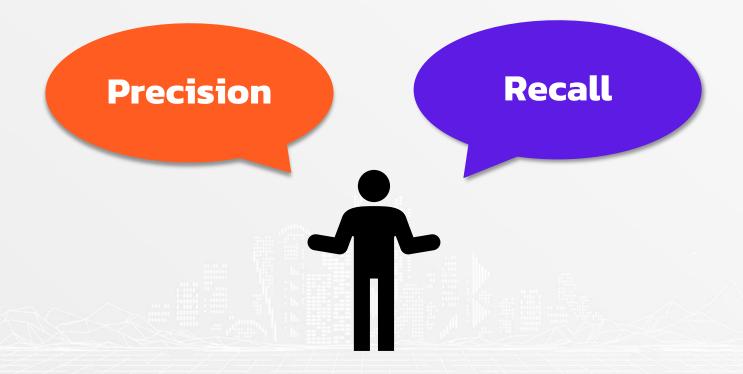


F1 Score

- What is F1 Score?
- Formula
- Step to calculate F1 Score
- Example
- Code
- F1 Score with imbalanced class



F1 Score คือ ค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall





Harmonic mean เป็นการหาค่าเฉลี่ยประเภทหนึ่ง โดยการหาค่าเฉลี่ยวิธีนี้จะ

- ให้น้ำหนักน้อยกับข้อมูลที่มีค่าเยอะ
- ให้น้ำหนักเยอะกับข้อมูลที่มีค่าน้อย



ให้**น้ำหนักเยอะ**กับข้อมูลที่มี**ค่าน้อย**

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

- ◆Harmonic Mean = 0.14939025281869237
- ◆ Arithmetic Mean = 32.667333333333333



Harmonic mean สามารถเขียนให้อยู่ในรูปดังต่อไปนี้

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

โดยที่ • n คือจำนวนข้อมูลทั้งหมด

 $lacktriangle x_i$ คือข้อมูลที่ i



ตัวอย่าง กำหนดให้ $\mathbf{x} = \{10, 100\}$

Harmonic mean จะสามารถคำนวณได้ดังต่อไปนี้

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



F1 Score จะพิจารณาค่าระหว่าง precision และ recall ผ่าน harmonic mean ซึ่ง สามารถจัดรูปได้ดังนี้

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



โดย F1 score จะสามารถตีความได้ดังต่อไปนี้

- ถ้าค่า F1 score มาก หมายความว่า ค่า precision และ recall มีค่ามากทั้งคู่
- ถ้าค่า F1 score <mark>น้อย</mark> หมายความว่า ค่า precision และ recall มีค่าน้อยทั้งคู่ หรือมี ค่าใดค่าหนึ่งน้อยมาก ๆ



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 \widehat{y}_i
- 2. หาค่า TP,FP และFN ของ model
- 3. หาค่า *precision* ของ model
- 4. หาค่า recall ของ model
- 5. วัดประสิทธิภาพของ model ตามสูตรของ F1



1. เก็บค่า y_i และ \hat{y}_i

	${oldsymbol{y_i}}$	$\widehat{\boldsymbol{y}}_{m{i}}$
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**



คำนวณค่า TP, FP และ FN

	y_i	\widehat{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)	o (ไม่เป็น โรค)
	lict	1 (เป็นโรค)	1 TP	fP 1
1 10 1 21 1 21 1 21 1 1 21 1 1 21 1 1 21	Predict	o (ไม่เป็น โรค)	2 FN	3 TN



3. หาค่า precision

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

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

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



4. หาค่า recall

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

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

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



5. วัดประสิทธิภาพของ model ตามสูตรของ F1

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

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

$$=\frac{2}{5}$$
$$=0.4$$



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)

3.4



F1 Score with imbalanced class

F1 Score สามารถใช้วัดประสิทธิภาพของ model ที่เป็น imbalanced class ได้ **ตัวอย่าง** ให้หมอปลอมตรวจโรคคนไข้จำนวน 100 คน <u>หมอปลอมตรวจพบว่าคนไข้ไม่</u> เป็นโรคเกือบทั้งหมด

		Actual	
		เป็นโรค	ไม่เป็นโรค
Predict	เป็นโรค	1 TP	OFP
	ไม่เป็นโรค	7 FN	92 TN



F1 Score with imbalanced class

• คำนวณผ่าน accuracy จะได้ค่าสูงถึง **0.93**

		Actual	
		เป็นโรค	ไม่เป็นโรค
Predict	เป็นโรค	1 TP	O
	ไม่เป็นโรค	7 FN	92 ^{TN}

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



F1 Score with imbalanced class

• คำนวณผ่าน **F1 score** จะได้ค่าออกมาเพียง **0.22**

		Actual	
		เป็นโรค	ไม่เป็นโรค
Predict	เป็นโรค	1 TP	O
	ไม่เป็นโรค	7 FN	92 ^{TN}

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}$$
$$= 2 \times \frac{1 \times \frac{1}{8}}{1 + \frac{1}{8}}$$



Model Evaluation in Classification

- **✓** Confusion Matrix
- Accuracy Score
- **✓• Precision Score**
- **✓• Recall Score**
- **√** F1 Score
 - Conclusion



Conclusion

Name	Formula	How to use	
Accuracy	$accuracy = \frac{TP + TN}{n}$	• ใช้กับการวัด model ที่ไม่เป็น imbalanced class	
Precision	$precision = \frac{TP}{TP + FP}$	 ใช้วัด model ที่ต้องการให้การเกิด False Positive น้อยๆ 	
Recall	$recall = \frac{TP}{TP + FN}$	• ใช้วัด model ที่ต้องการให้เกิด False Negative น้อยๆ	
F1	$F1 = 2 \times \frac{precision \times recall}{precision + recall}$	 ใช้วัด model ที่เป็น imbalanced class ได้ ต้องการพิจารณาทั้ง precision และ recall พร้อมกัน 	



Model Evaluation in Classification

- **✓** Confusion Matrix
- Accuracy Score
- **✓• Precision Score**
- **✓• Recall Score**
- **√** F1 Score
- **✓** Conclusion