

# Machine Learning



*2021-01-26*

*Science Team 17기 신범교*

# OUTLINES

①

## 1. Machine Learning 개요

②

머신러닝이란?

③

머신러닝 vs 딥러닝

머신러닝 유형 (지도학습, 비지도학습, 강화학습)

## 2. 지도학습

지도학습의 일반적인 프로세스

Decision Tree (의사결정나무)

## 3. 머신러닝 모델 평가

## 4. 실습



①

②

③

# 1. Machine Learning 개요

# 1. Machine Learning 개요

①

## Machine Learning (기계학습) 이란?

②

③



인공지능의 한 분야로, 컴퓨터가 데이터를 스스로 학습하여  
특정 목적(예측-분류, 회귀 / 해석 등)을 수행하는 것

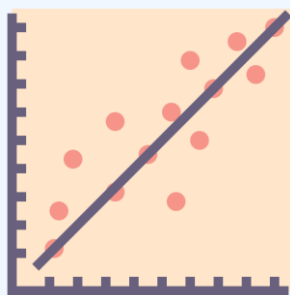
# 1. Machine Learning 개요

①

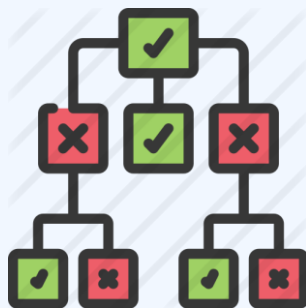
## Machine Learning vs Deep Learning

②

### Machine Learning



Linear Regression

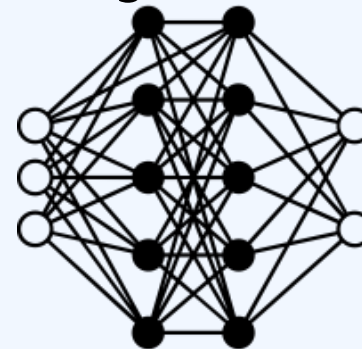


Decision Tree

- 다양한 구조의 학습 방법이 있음 (선형, Tree형 등)
- 기본 모델의 형태가 비교적 간단해 해석이 용이함
- 인간이 데이터를 모델 형태에 맞게 가공해서 제공해야 함 (정규분포 등의 통계적 가정)

고전적인 정형 데이터(데이터 프레임 형태) 위주의 학습

### Deep Learning



- 인간의 신경망을 본따 만든 구조의 학습 방법
- 신경망이 복잡하게 얽혀 있어 해석이 어려움 (블랙박스)
- 데이터에 대한 통계적 가정이 크게 필요 없음
- 더 많은 데이터와 높은 차원의 연산을 요구/가능

비정형 데이터(이미지, 텍스트 등)를 다루기에 용이함

# 1. Machine Learning 개요

## Machine Learning 의 유형



# 1. Machine Learning 개요

## 지도 학습(Supervised Learning) vs 비지도 학습 (Unsupervised Learning)

학습 유형	행위
지도 학습	과거의 판매 데이터(X)를 기반으로 학습해 특정 요일의 매출(Y) 예측 회귀 (Regression)
	기존 의류 샘플 이미지(X)를 학습해 신상 의류 카테고리(Y) 분류 분류 (Classification)
비지도 학습	고객의 체형 별 구매 패턴을 분석하기 위해 몸무게(X1)와 키(X2)가 비슷한 그룹 생성 군집화 (Clustering)
	고객이 스웨터를 사면(X1) 높은 확률로 셔츠를 산다(X2) 등의 규칙 발견 연관 규칙 (Association Rule)

지도 학습과 비지도 학습의 간단한 구분법은 '학습의 결과에 대해 정답이 있는가?'



①

②

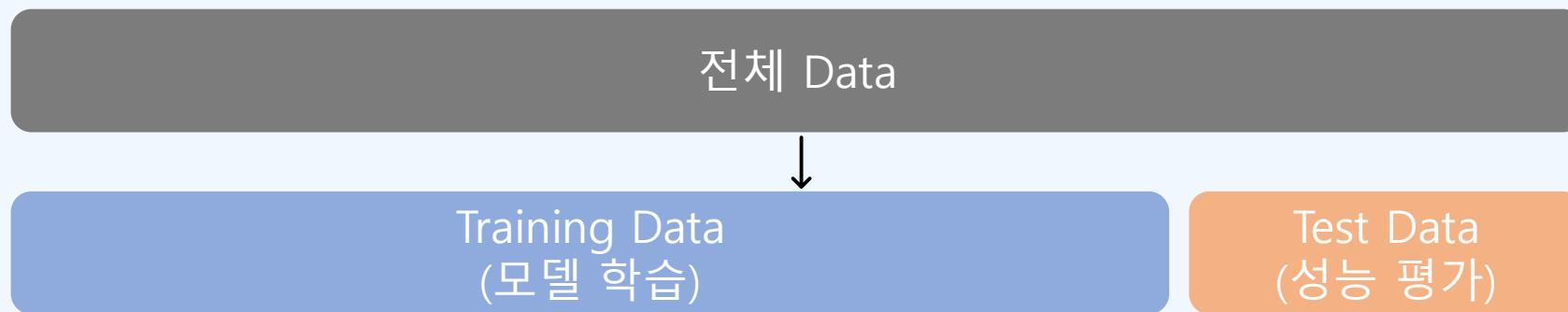
③

## 2. 지도 학습 (Supervised Learning)



## 2. 지도 학습 (Supervised Learning)

### 지도 학습(Supervised Learning)의 일반적인 Process



1. 전체 Data를 Training / Test Data Set으로 분할 (이 때의 비율은 흔히 8:2, 7:3)
2. Training Data로 모델을 학습
3. Test Data의 실제 값( $Y$ ) 과 Test Data의 설명변수( $X$ )들을 통해 모델이 예측한 값 ( $\hat{Y}$ )과 비교하여 모델의 성능을 평가

## 2. 지도 학습 (Supervised Learning)

### 지도 학습(Supervised Learning)의 일반적인 Process

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	
0	1	0	3	Braund, Mr. Owen Harris	0	22.0	1	0	A/5 21171	7.2500	0.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	38.0	1	0	PC 17599	71.2833	1.0
2	3	1	3	Heikkinen, Miss. Laina	1	26.0	0	0	STON/O2. 3101282	7.9250	0.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	35.0	1	0	113803	53.1000	0.0
4	5	0	3	Allen, Mr. William Henry	0	35.0	0	0	373450	8.0500	0.0
5	6	0	3	Moran, Mr. James	0	30.0	0	0	330877	8.4583	2.0
6	7	0	1	McCarthy, Mr. Timothy J	0	54.0	0	0	17463	51.8625	0.0
7	8	0	3	Palsson, Master. Gosta Leonard	0	2.0	3	1	349909	21.0750	0.0
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	1	27.0	0	2	347742	11.1333	0.0
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	1	14.0	1	0	237736	30.0708	1.0

Q. 다음 Data Set에 대해 각 사람이 Survived 할지 분류할 수 있는 모델을 만드세요.

## 2. 지도 학습 (Supervised Learning)

### 지도 학습(Supervised Learning)의 일반적인 Process

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	
0	1	0	3	Braund, Mr. Owen Harris	0	22.0	1	0	A/5 21171	7.2500	0.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	38.0	1	0	PC 17599	71.2833	1.0
2	3	1	3	Heikkinen, Miss. Laina	1	26.0	0	0	STON/O2. 3101282	7.9250	0.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	35.0	1	0	113803	53.1000	0.0
4	5	0	3	Allen, Mr. William Henry	0	35.0	0	0	373450	8.0500	0.0
5	6	0	3	Moran, Mr. James	0	30.0	0	0	330877	8.4583	2.0
6	7	0	1	McCarthy, Mr. Timothy J	0	54.0	0	0	17463	51.8625	0.0
7	8	0	3	Palsson, Master. Gosta Leonard	0	2.0	3	1	349909	21.0750	0.0
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	1	27.0	0	2	347742	11.1333	0.0
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	1	14.0	1	0	237736	30.0708	1.0

종속변수(Y), Label, Target  
\*Survival 0=No, 1=Yes

설명변수(X), 독립변수, Feature, Input

## 2. 지도 학습 (Supervised Learning)

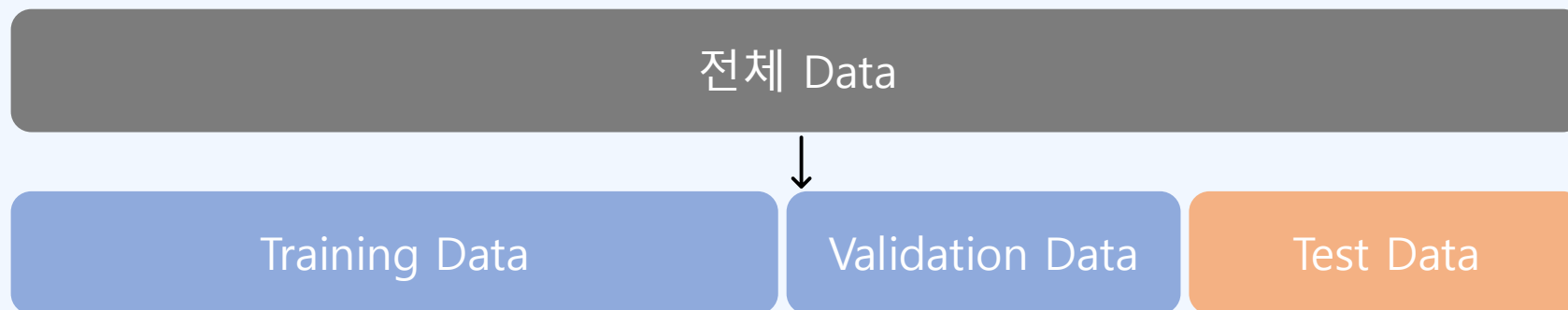
### 지도 학습(Supervised Learning)의 일반적인 Process

Training	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked		
	0	1	0	3	Braund, Mr. Owen Harris	0	22.0	1	0	A/5 21171	7.2500	0.0	
	1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	38.0	1	0	PC 17599	71.2833	1.0	
	2	3	1	3	Heikkinen, Miss. Laina	1	26.0	0	0	STON/O2. 3101282	7.9250	0.0	
	3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	35.0	1	0	113803	53.1000	0.0	
	4	5	0	3	Allen, Mr. William Henry	0	35.0	0	0	373450	8.0500	0.0	
	5	6	0	3	Moran, Mr. James	0	30.0	0	0	330877	8.4583	2.0	
	6	7	0	1	McCarthy, Mr. Timothy J	0	54.0	0	0	17463	51.8625	0.0	
	Test	7	8	0	3	Palsson, Master. Gosta Leonard	0	2.0	3	1	349909	21.0750	0.0
		8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	1	27.0	0	2	347742	11.1333	0.0
9		10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	1	14.0	1	0	237736	30.0708	1.0	
종속변수(Y)			설명변수(X)										

1. Training Data의 종속변수(Y)와 설명변수(X)로 모델을 학습
2. Test Data의 설명변수(X)를 모델에 입력해 예측 값( $\hat{Y}$ ) 출력
3. Test Data의 실제 종속변수(Y)와 예측 값( $\hat{Y}$ )를 비교해 모델 성능 평가

## 2. 지도 학습 (Supervised Learning)

지도 학습(Supervised Learning)의 일반적인 Process

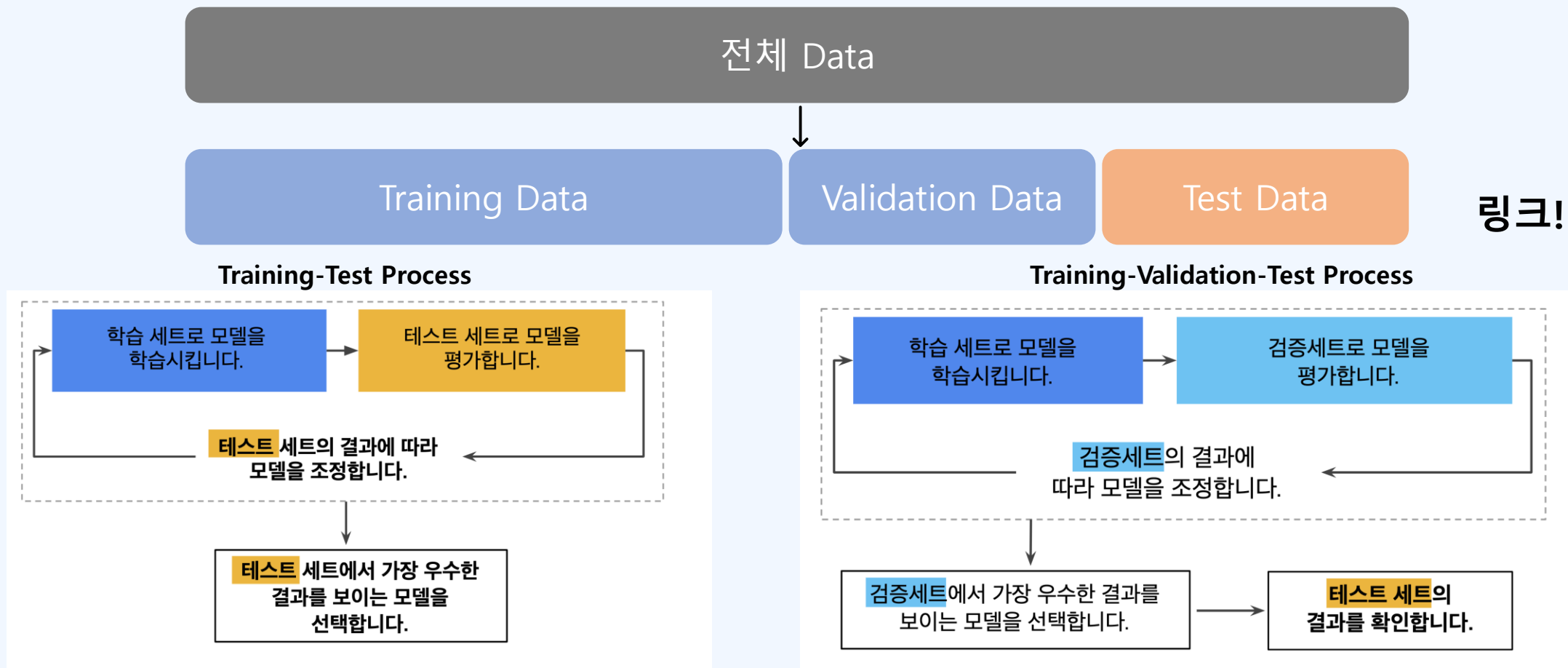


**Validation Data**는 학습 과정에서 모델의 성능을 검증하기 위해 사용하는 데이터  
**Test Data**는 모델의 최종 성능을 평가하는데 사용

**물론, Validation Data가 없어도 모델링은 가능하다!**  
Validation Data를 활용하는 것은 더 안정적인 모델을 만들기 위해서

## 2. 지도 학습 (Supervised Learning)

### 지도 학습(Supervised Learning)의 일반적인 Process



## 2. 지도 학습 (Supervised Learning)

①

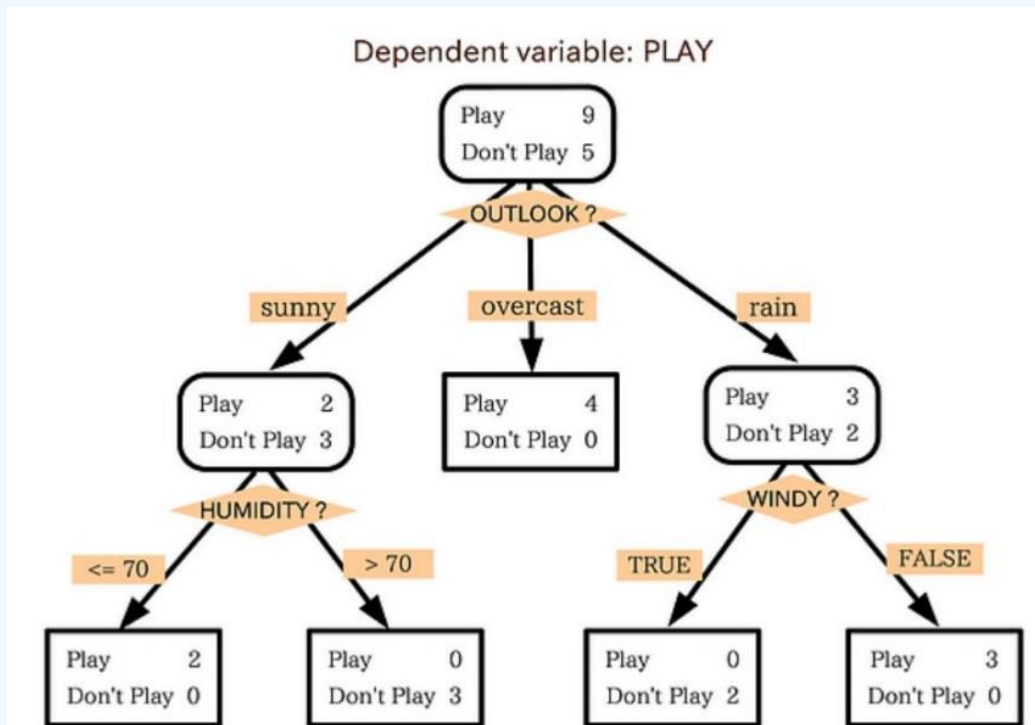
### Decision Tree (의사결정나무)

②

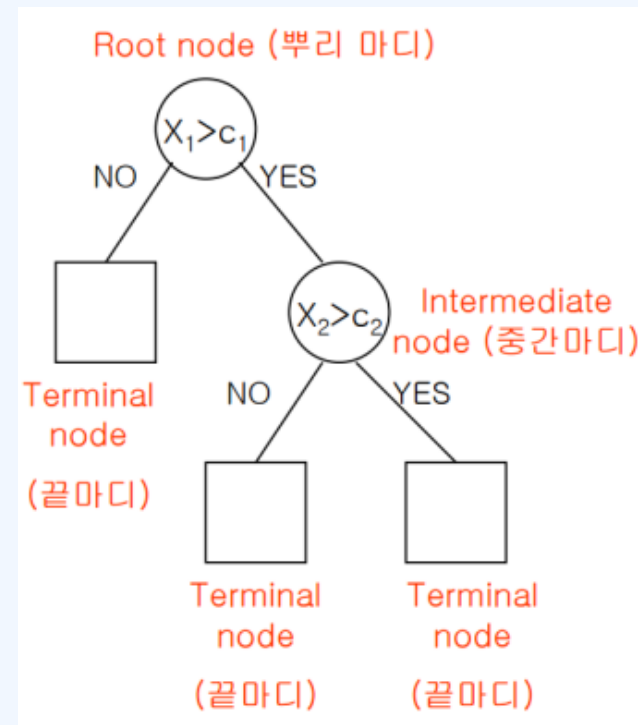
데이터 사이에 존재하는 패턴을 규칙들의 조합으로 나타내며, 이를 그림으로 나타내면 그 구조가 마치 나무와 같다고 해서 Decision Tree(의사결정나무)라 불린다.

\* 질문을 던져서 대상을 좁혀 나가는 '스무고개' 놀이와 비슷한 개념

③



Decision Tree 예시



일반화 된 Decision Tree 구조

## 2. 지도 학습 (Supervised Learning)

①

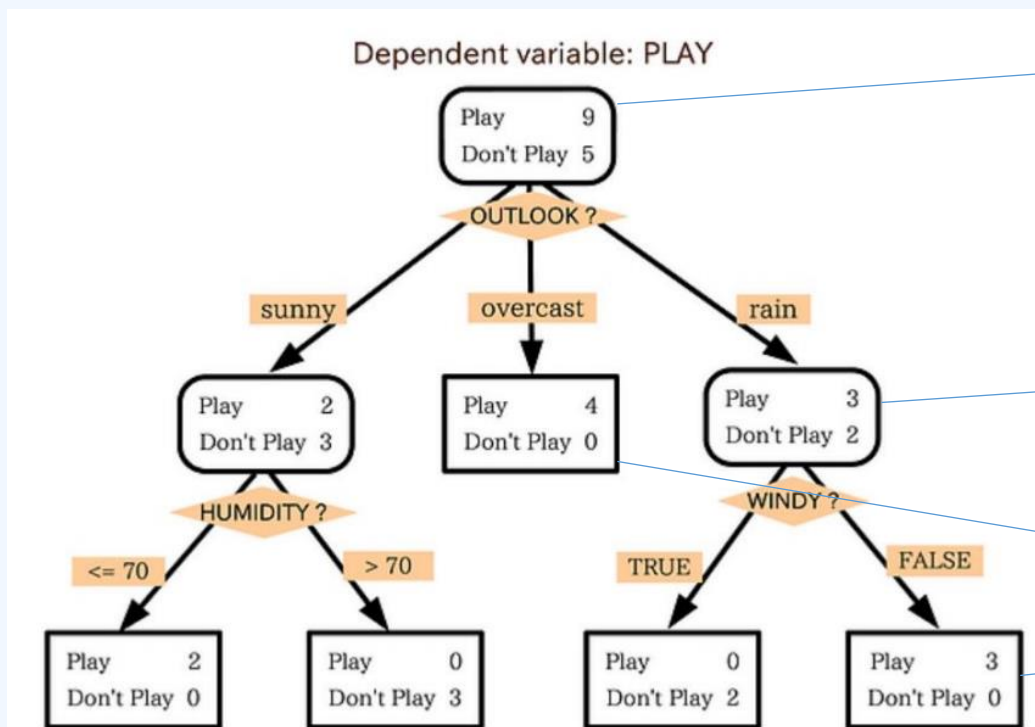
### Decision Tree (의사결정나무)

②

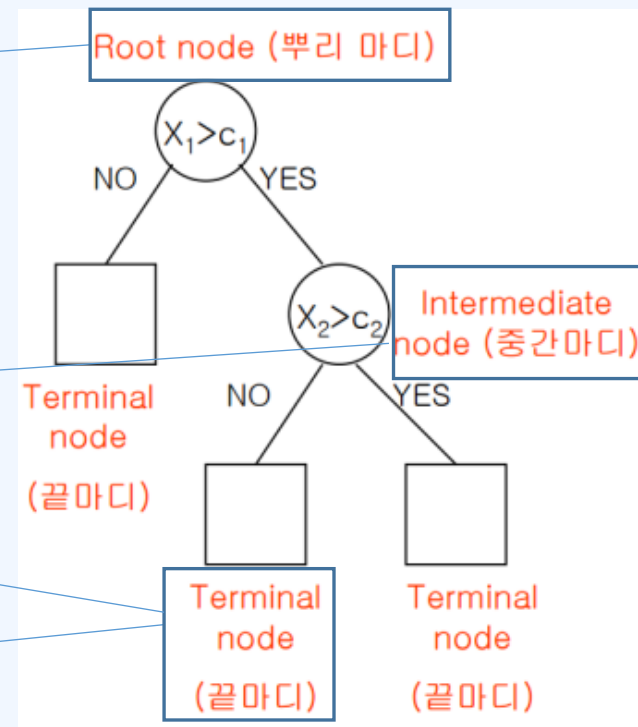
데이터 사이에 존재하는 패턴을 규칙들의 조합으로 나타내며, 이를 그림으로 나타내면 그 구조가 마치 나무와 같다고 해서 Decision Tree(의사결정나무)라 불린다.

\* 질문을 던져서 대상을 좁혀 나가는 '스무고개' 놀이와 비슷한 개념

③



Decision Tree 예시



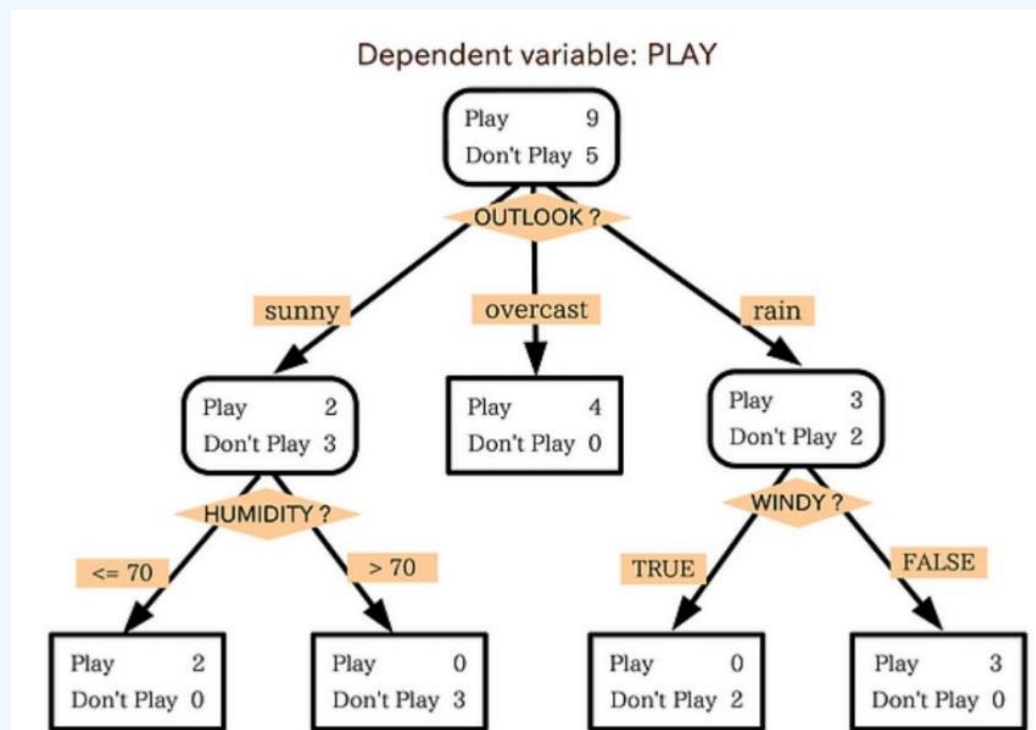
일반화 된 Decision Tree 구조



## 2. 지도 학습 (Supervised Learning)

### Decision Tree (의사결정나무)

'운동경기가 열릴지 열리지 않을지 예측' 하기 위한 Decision Tree 모델

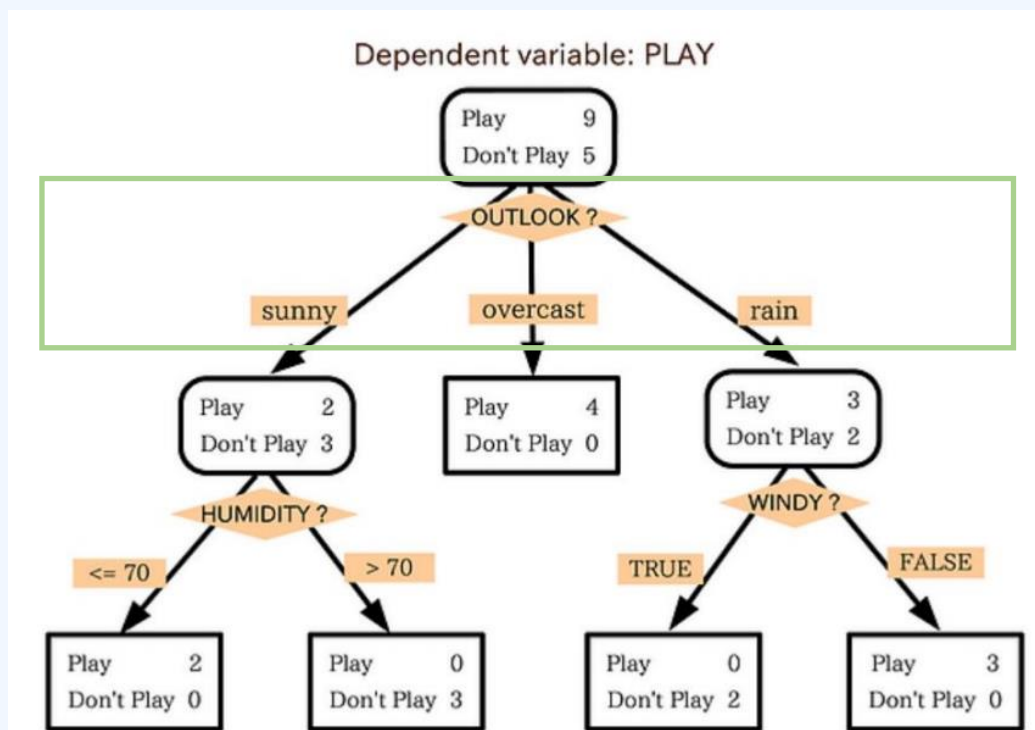


Y : 경기 진행/취소 여부  
X1 : 날씨 (OUTLOOK)  
X2 : 습도 (HUMIDITY)  
X3 : 바람 (WINDY)

## 2. 지도 학습 (Supervised Learning)

### Decision Tree (의사결정나무)

'운동경기가 열릴지 열리지 않을지 예측' 하기 위한 Decision Tree 모델



Y : 경기 진행/취소 여부

X1 : 날씨 (OUTLOOK)

X2 : 습도 (HUMIDITY)

X3 : 바람 (WINDY)

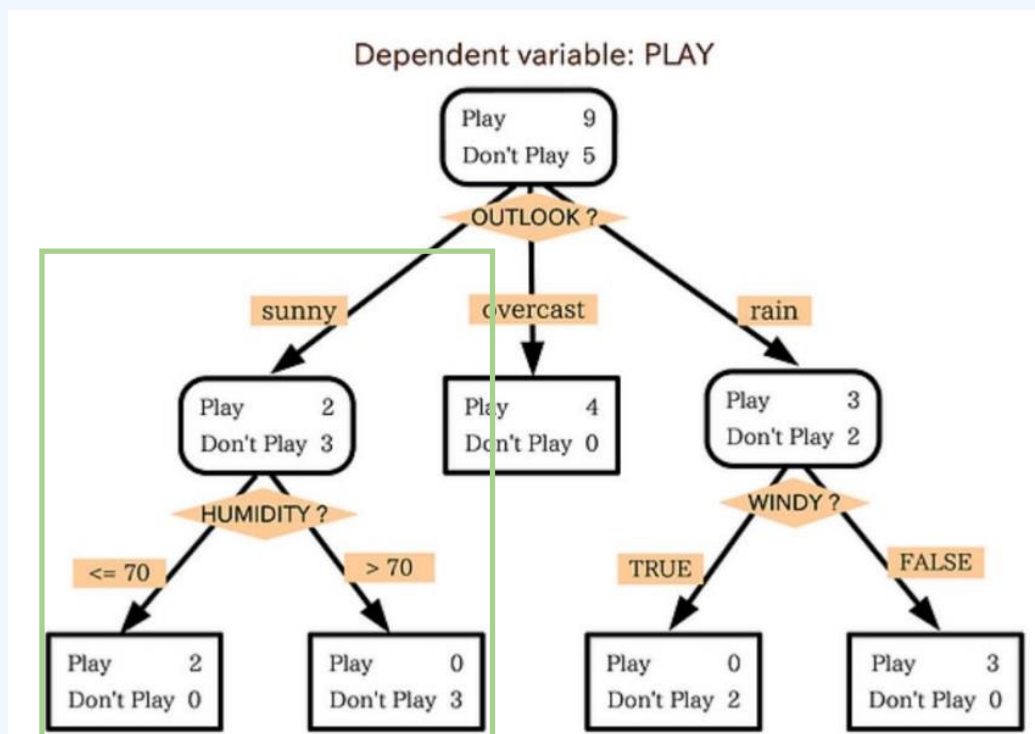
날씨	경기 진행 여부
맑음 (sunny)	경기 진행 3회, 경기 취소 2회
흐림 (overcast)	경기 진행 4회
비 (rain)	경기 진행 3회, 경기 취소 2회

흐린 날에는 경기 취소가 없으므로 분류 종료

## 2. 지도 학습 (Supervised Learning)

### Decision Tree (의사결정나무)

'운동경기가 열릴지 열리지 않을지 예측' 하기 위한 Decision Tree 모델



Y : 경기 진행/취소 여부

X1 : 날씨 (OUTLOOK)

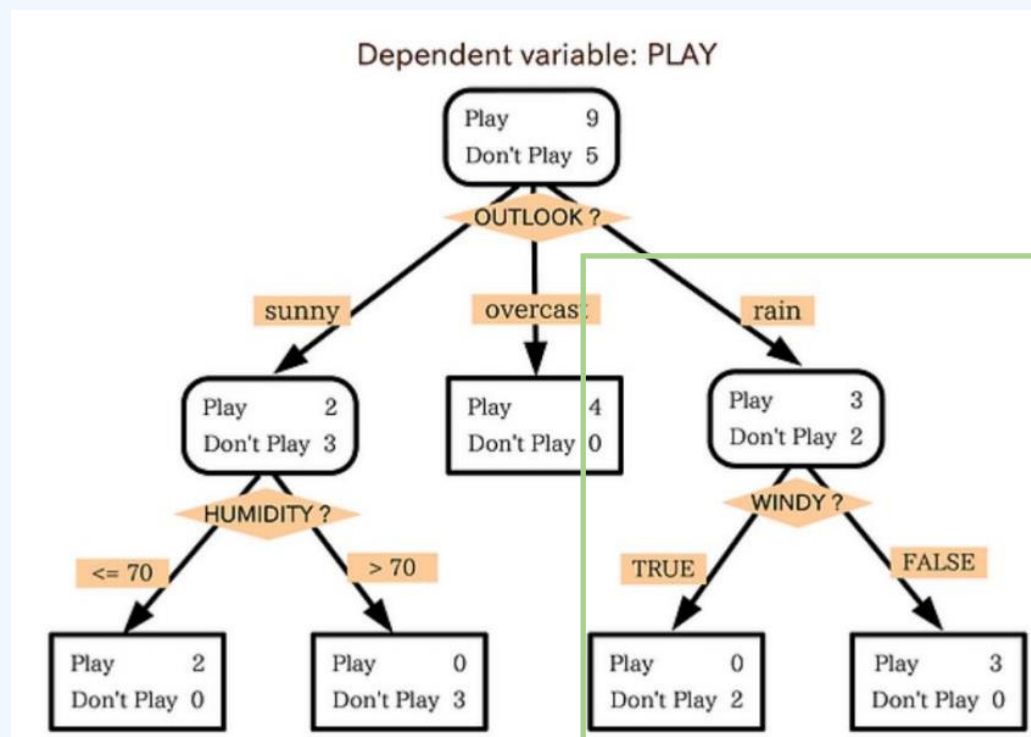
**X2 : 습도 (HUMIDITY)**

X3 : 바람 (WINDY)

## 2. 지도 학습 (Supervised Learning)

### Decision Tree (의사결정나무)

'운동경기가 열릴지 열리지 않을지 예측' 하기 위한 Decision Tree 모델



Y : 경기 진행/취소 여부

X1 : 날씨 (OUTLOOK)

X2 : 습도 (HUMIDITY)

X3 : 바람 (WINDY)

## 2. 지도 학습 (*Supervised Learning*)

①

### Decision Tree의 특징

②

1. 분류(Classification)와 회귀(Regression)가 모두 가능

③

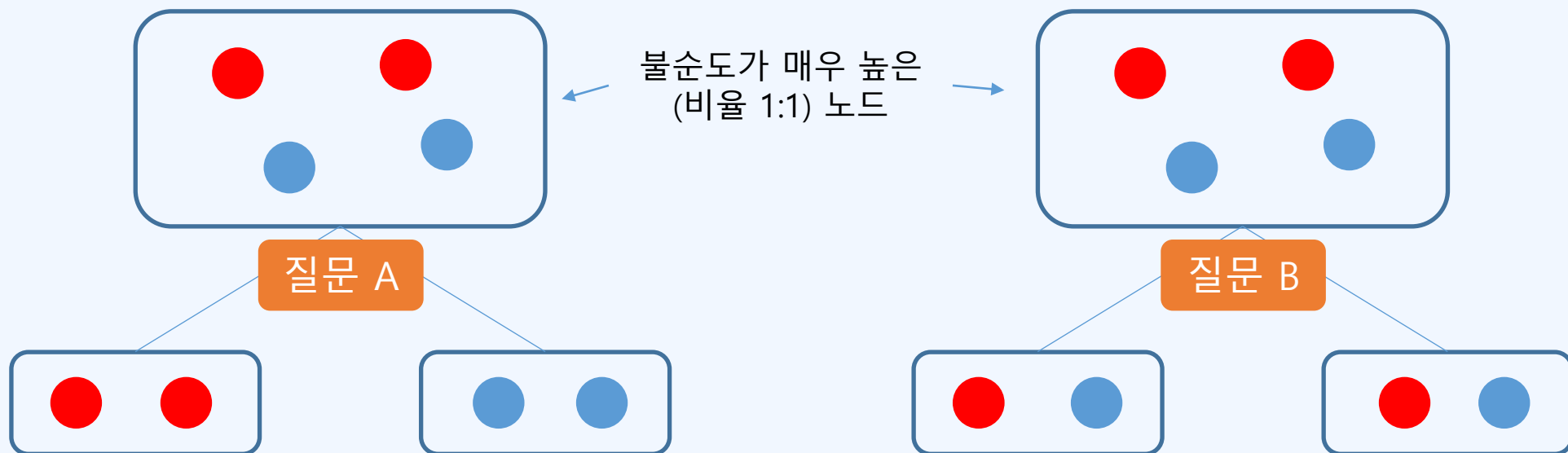
2. 불순도(Impurity)를 줄이는 방향으로 학습 진행

3. 다양한 **Ensemble** 기법의 기초가 되는 알고리즘

## 2. 지도 학습 (Supervised Learning)

질문을 던지는 기준: 불순도(Impurity)와 Information Gain

불순도(Impurity)는 종류가 다른 데이터가 여러 개 섞여 있음을 의미  
이를 줄이는 방향으로 학습을 진행한다 -> 분류를 제일 잘 시키는 질문을 찾는다



이렇게 불순도를 줄이게 되는 것을 **Information Gain** 이라고 표현하고,  
분류 전의 불순도에서 분류 후의 불순도를 빼서 구한다.  
분류 시에 **Information Gain**을 높게 만드는 질문을 선택

## 2. 지도 학습 (Supervised Learning)

### 불순도 지표와 Decision Tree 알고리즘 종류

불순도(Imputiry)는 종류가 다른 데이터가 여러 개 섞여 있음을 의미  
이를 줄이는 방향으로 학습을 진행한다 -> 분류를 제일 잘 시키는 질문을 찾는다

알고리즘	불순도 지표(분류 기준)	비고
CART (Classification and Regression Tree)	범주형 -> Gini Index	종속변수가 범주형, 이산형 모두 사용 가능
	연속형 -> 분산의 감소량	
CHAID	카이제곱 통계량	종속변수가 이산형 일 때 사용  과적합(Overfitting) 전에 스스로 분류 중지
ID3, C4.5, C5.0	Entropy	종속변수가 범주형 일 때 사용

Python 등 대부분의 툴에서 기본적으로 구현된 Decision Tree 알고리즘은 CART를 이용한 것

## 2. 지도 학습 (Supervised Learning)

①

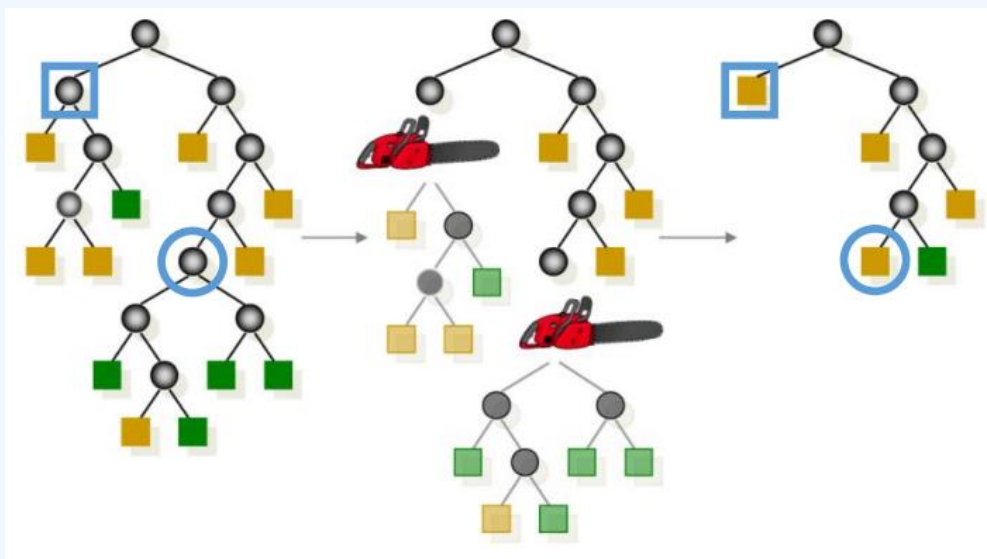
**Q. Decision Tree를 설계할 때, 모든 데이터가 다 분류될 때 까지 가지를 치는가?**

②

Training Data에 너무 과도하게 적합 되어 새로운 데이터를 올바르게 예측하기 어렵다.

③

Decision Tree는 Pruning이라는 과정을 통해 과도한 적합을 방지하고 일반성을 확보.  
Python의 Scikit-Learn에서는 Pre-Pruning을 Hyper Parameter를 통해 설정할 수 있다.



Pruning 유형

방법 예시

Pre-Pruning

최대 깊이 제한,  
노드 수 제한,  
feature 수 제한 등

Post-Pruning

사용자가 지정한 기  
준에 맞춰 모델 학습  
후 가지 제거



## 2. 지도 학습 (Supervised Learning)

### Decision Tree 요약

#### 장점

- 인과관계 설명(Interpretation)이 쉽다
- 인간의 의사결정 과정과 닮았다
- 범주형 자료를 변형 없이 다룰 수 있다
- 이상치(Outlier, Noise)에 강하다
- 통계적 가정이 크게 필요하지 않다

#### 단점

- 일반적으로 단독 예측력(Prediction)이 뛰어나지 않다
- 과적합에 취약하다
- Training Data가 조금만 바뀌어도 Tree 모양이 크게 바뀐다



①

②

③

## 3. 머신러닝 모델 평가

### 3. 머신러닝 모델 평가

①

모델의 성능을 어떻게 평가(Evaluation) 할 것인가?

②

지도 학습의 목표는 분류/회귀/예측 이므로, 우리가 만들어낸 머신러닝 모델이 과연 올바르게 목표를 수행하고 있는냐에 대한 평가 기준이 필요

③

회귀(Regression)-연속형 자료	분류(Classification)-범주형 자료
Mean Squared Error(MSE)	Accuracy
$R^2$ (결정계수)	$F_1$
Adjusted $R^2$	ROC-AUR
...	...

#### Binary Classification

예측을 Yes / No 두 가지로만 진행하는 분류 문제

### 3. 머신러닝 모델 평가

모델의 성능을 어떻게 평가(Evaluation) 할 것인가?

TransactionID	IsFraud	TransactionDT	TransactionAmt	ProductCD	card1	card2	card3	card4	card5	...
3059884	0	1623818	23.087	C	15885	545.0	185.0	visa	138.0	...
3375522	0	9747360	37.340	C	14675	368.0	212.0	visa	226.0	...
3193413	1	4743953	29.000	W	11233	321.0	150.0	visa	195.0	...
3063585	0	1686148	250.000	R	6019	583.0	150.0	visa	226.0	...
3234787	0	5885169	300.000	W	2616	321.0	150.0	discover	102.0	...
3074481	0	1837414	11.194	C	3154	408.0	185.0	mastercard	224.0	...
3059940	0	1624327	100.000	R	6019	583.0	150.0	visa	226.0	...
3002100	1	419140	40.000	W	3507	321.0	150.0	visa	226.0	...
3234741	0	5885169	300.340	W	2616	321.0	150.0	discover	102.0	...
3073154	1	1837414	11.194	C	3154	408.0	185.0	visa	224.0	...

Q. 다음 거래 정보에 대한 Data Set에 대해 각 거래가 사기 거래인지 정상 거래인지 분류할 수 있는 모델을 만드세요

### 3. 머신러닝 모델 평가

#### Confusion Matrix (오차행렬/ 혼동행렬)

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

예측에 대한 표현은  
Positive / Negative

예측 성공 여부는  
True / False

### 3. 머신러닝 모델 평가

#### Accuracy (정확도)

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	47	4
Not Fraud (0)	3	46

Accuracy

$$= \frac{\text{True Positive} + \text{True Negative}}{\text{Total OBS}}$$

총 관측치 (100건) 중  
올바르게 예측한 비율

$$\frac{47 + 46}{100} = 0.93 \quad 93\% \text{ Accuracy}$$

### 3. 머신러닝 모델 평가

#### Accuracy (정확도) 의 한계점

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total OBS}}$$

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)		
Not Fraud (0)		

**Fraud**  
3건

**Not Fraud**  
97건

### 3. 머신러닝 모델 평가

①

#### Accuracy (정확도) 의 한계점

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total OBS}}$$

②

③

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)		
Not Fraud (0)		

**Fraud**  
3건

**Not Fraud**  
97건

**사기 예측**  
0건

**정상 예측**  
100건



### 3. 머신러닝 모델 평가

①

#### Accuracy (정확도) 의 한계점

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total OBS}}$$

②

③

실제값 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	0	0
Not Fraud (0)	3	97

**Fraud**  
3건

**Not Fraud**  
97건

**사기 예측**  
0건

$$\frac{97 + 0}{100} = 0.97$$

97% Accuracy?

**정상 예측**  
100건

**Target 값의 실제 분포가  
Imbalance 한 경우 신뢰  
도가 떨어진다**

### 3. 머신러닝 모델 평가

$F_1$ , ROC-AUC :: Recall, Precision, Specificity

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

### 3. 머신러닝 모델 평가

$F_1$ , ROC-AUC :: Recall, Precision, Specificity

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

실제 Fraud 를 올바르게 가려낸 비율

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

### 3. 머신러닝 모델 평가

$F_1$ , ROC-AUC :: Recall, Precision, Specificity

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

실제 Not Fraud 를 올바르게 가려낸 비율

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

### 3. 머신러닝 모델 평가

$F_1$ , ROC-AUC :: Recall, Precision, Specificity

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

예측 Fraud 중 실제 Fraud 를 올바르게 예측한 비율

### 3. 머신러닝 모델 평가

$F_1$ , ROC-AUC :: Recall, Precision, Specificity

예측값 \ 실제값	Fraud (1)	Not Fraud (0)
Fraud (1)	0	0
Not Fraud (0)	3	97

$$\text{Recall} = \frac{TP}{TP+FN} = 0$$

$$\text{Precision} = \frac{TP}{TP+FP} = 0$$

$$\text{Specificity} = \frac{TN}{TN+FP} = 0$$

### 3. 머신러닝 모델 평가

#### $F_1$ Score

Recall과 Precision을 동시에 반영할 수 있도록 고안된 지표

실제값 \ 예측값	Fraud (1)	Not Fraud (0)
Fraud (1)	TP	FP
Not Fraud (0)	FN	TN

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

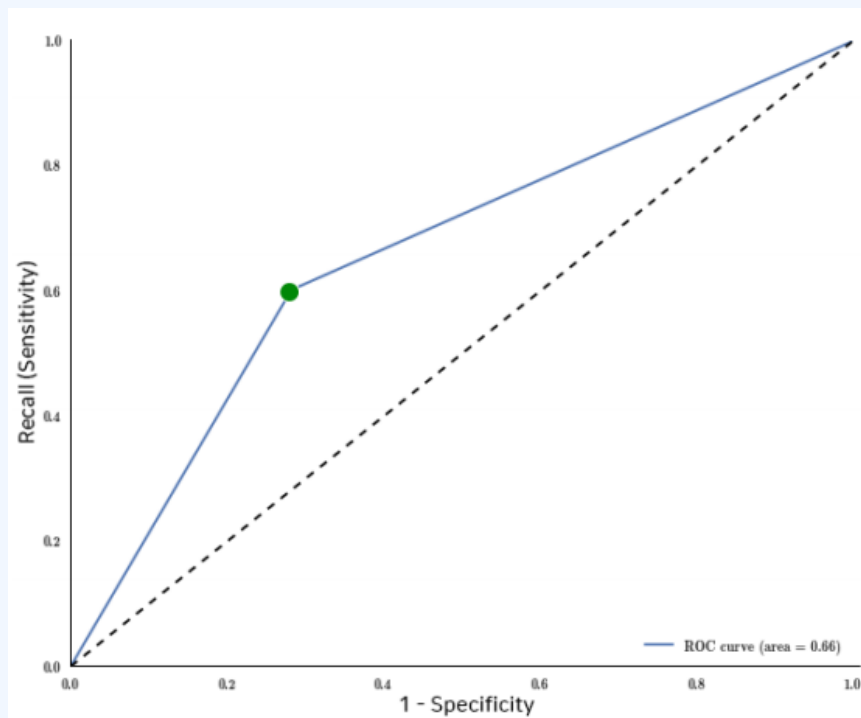
$$F_1 \text{ Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}$$

[0,1] 값을 가지며, 1에 가까울 수록 잘 분류되었다고 평가

### 3. 머신러닝 모델 평가

#### ROC-AUC

ROC Graph



(1-Specificity, Recall) 값을 좌표평면에 도식하고  
그래프 원점, 종점과 연결한 Graph

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

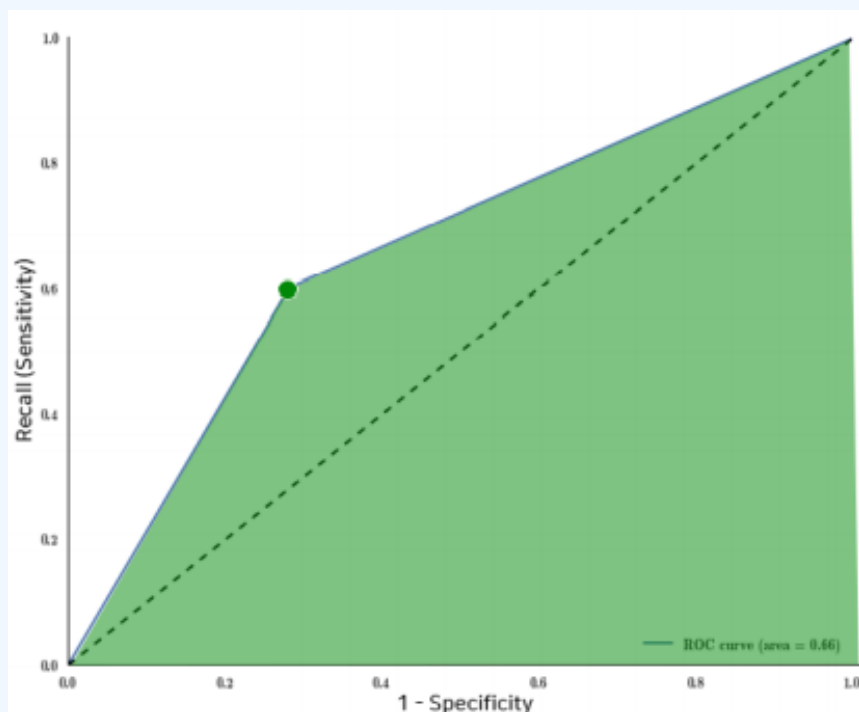
$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$



### 3. 머신러닝 모델 평가

#### ROC-AUC

ROC Graph



(1-Specificity, Recall) 값을 좌표평면에 도식하고  
그래프 원점, 종점과 연결한 Graph

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

#### ROC-AUC

#### ROC Graph 밑 면적의 넓이

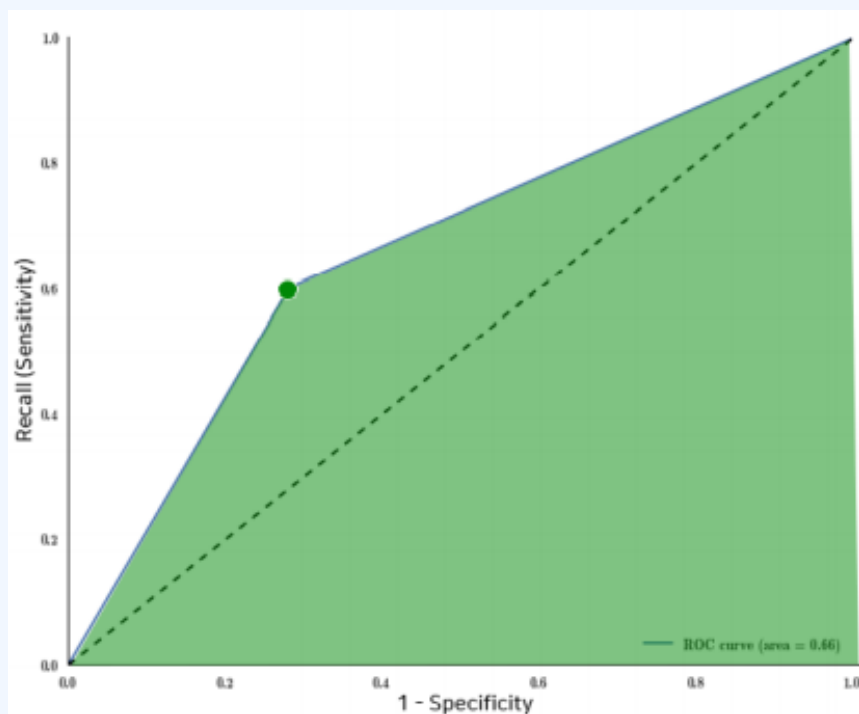
[0,1] 값을 가지며, Recall과 Specificity를 고려한 지표  
1에 가까울수록 잘 분류되었다고 평가

### 3. 머신러닝 모델 평가

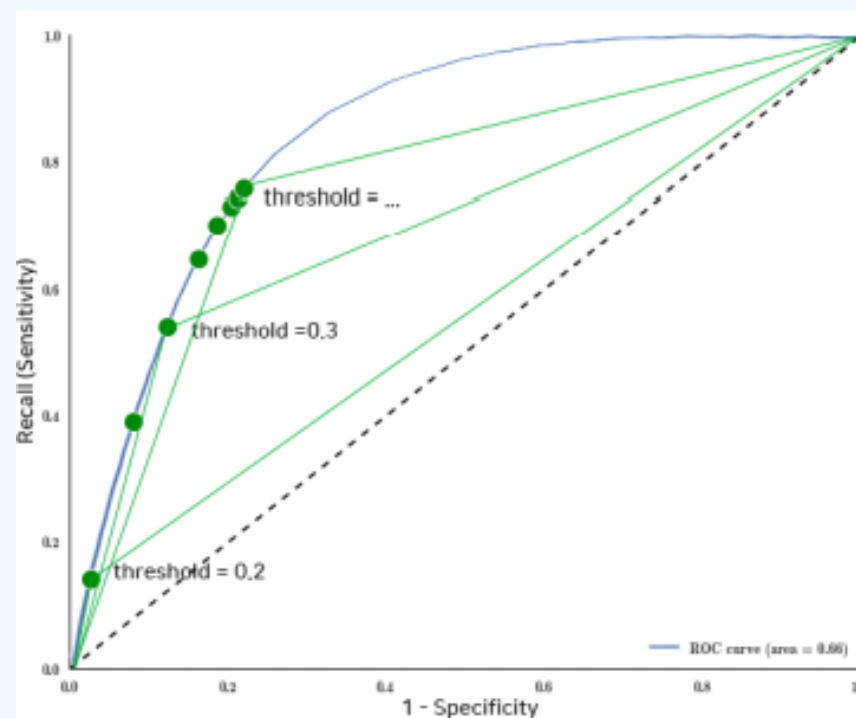
링크

#### ROC-AUC

ROC Curve :: 몇몇 모델의 Output은 **Target에 대한 확률** (ex. Logistic Regression)



Threshold 선언 시  
(1,0 binary 형태의 output)



Threshold 미 선언 시  
(Probability 형태의 output)