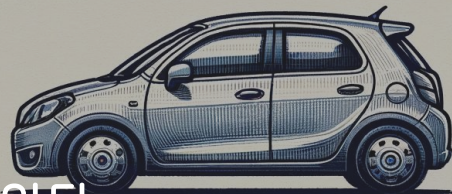
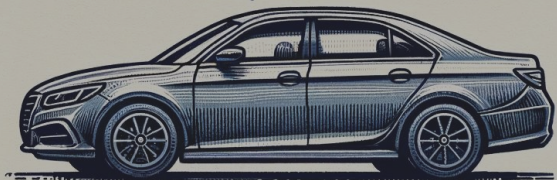


# 주차장 규칙

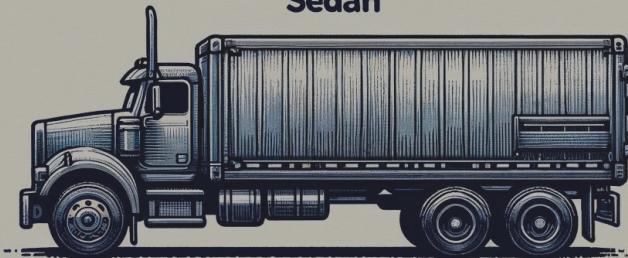
- 경차, 승용차, 트럭 세 가지 종류의 차를 주차할 수 있다
- 각각 차 종류에 맞는 주차 자리가 있다
- 경차는 경차 자리에,
- 승용차는 승용차 자리에,
- 트럭은 트럭 자리에 주차를 해야 한다



Compact Car



Sedan



Truck





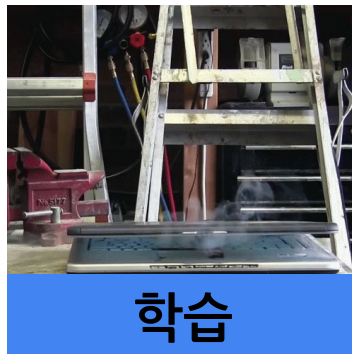
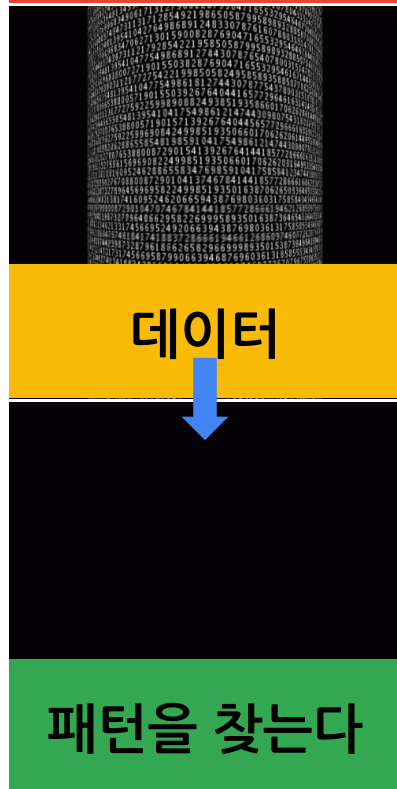
# 전통적인 프로그래밍 방식



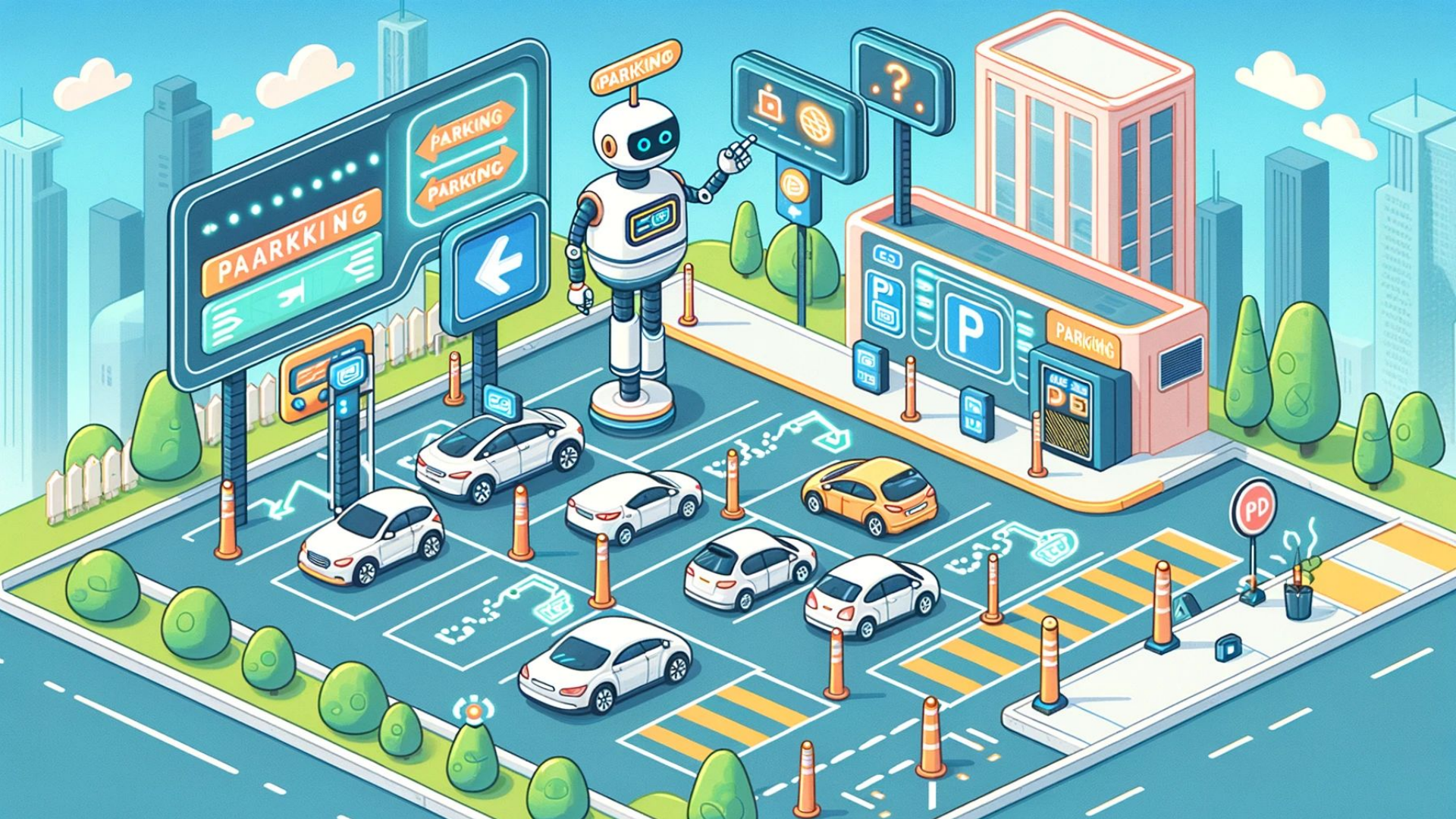




# 머신러닝을 이용한 프로그래밍 방식









DATA

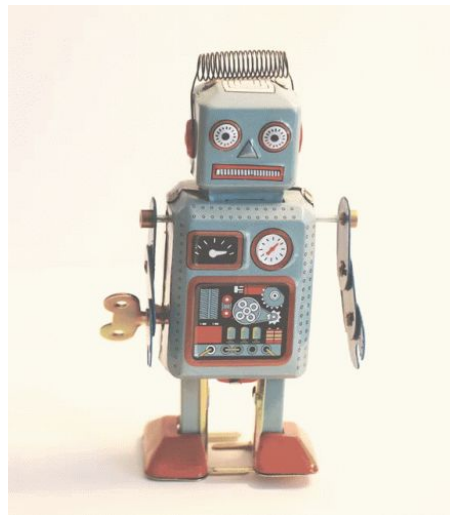
LEARNING

MODEL





# 데이터의 중요성



garbage in, garbage out



# 인공지능 (AI)

## 머신러닝

## 딥러닝

————— 기계가 인간처럼 지능적으로 작동하게 하는 기술

————— 데이터로부터 학습하여 예측을 수행하는 AI의 한 방법

————— 머신러닝의 한 기술로, 복잡한 문제를 해결하기 위해 **깊은 신경망**을 사용





# 알고리즘

머신러닝에서 알고리즘은 데이터에서 패턴을 학습하는 방법

레시피에 따라 요리를 하는 것처럼, 알고리즘은 데이터를 사용해 특정 작업(예: 이메일이 스팸인지 아닌지 판별하기)을 수행하는 방법을 컴퓨터에 알려줍니다.



# 모델 (Model)

머신러닝 모델은 알고리즘을 통해 데이터로부터 학습된 지식의 집합입니다.

데이터를 분석한 후, 알고리즘이 추출한 패턴을 기억하는데, 이렇게 학습된 정보를 모델이라고 합니다.

이 모델을 사용하여 새로운 데이터에 대한 예측을 할 수 있습니다.

모델은 머신러닝에서 뇌와 같은 역할을 합니다.





# 훈련 (Training)

훈련은 모델이 데이터로부터 학습하는 과정을 말합니다.

학교에서 공부하여 시험을 치르는 것처럼, 머신러닝 모델도 대량의 데이터를 '공부'하여 어떤 패턴이 있는지를 '학습'합니다.

이 과정에서 모델은 데이터 속 숨어 있는 규칙이나 관계를 찾아내어 이를 기억합니다.



# 예측 (Prediction)

예측은 모델이 새로운 데이터에 대해 결론을 내리는 것을 말합니다.

훈련을 통해 충분히 학습한 모델은 이전에 본 적 없는 새로운 데이터에 대해서도 학습한 지식을 바탕으로 무엇인지를 '예측'할 수 있습니다.

예를 들어, 과거의 집값 데이터를 통해 학습한 모델은 새로운 집에 대한 가격을 예측할 수 있습니다.





# 데이터 세트 (Dataset)

데이터 세트는 머신러닝 알고리즘이 학습하는 데 사용되는 데이터의 집합입니다.

마치 시험 공부를 위한 교과서처럼, 데이터 세트는 모델이 학습하는 데 필요한 예제와 정보를 담고 있습니다.



# 특성 (Feature)

특성은 데이터의 개별적인 관측치나 측정치를 말합니다.

예를 들어, 집을 설명할 때 방의 개수나 위치, 면적 등이 특성이 됩니다.

머신러닝에서는 이러한 특성들을 사용하여 패턴을 학습합니다.



# 레이블 (Label)

지도 학습에서, 레이블은 각 데이터 포인트에 대한 정답 또는 결과값을 말합니다.

예를 들어, 사진이 고양이인지 개인지를 나타내는 태그와 같습니다.





# 지도 학습 (Supervised Learning)

지도 학습은 레이블이 있는 데이터를 사용하여 모델을 훈련시키는 방식입니다.

여기서 '지도'는 교사가 정답을 가르쳐 주는 것처럼, 모델에게 올바른 예측을 '가르치는' 과정을 의미합니다.



# 비지도 학습 (Unsupervised Learning)

비지도 학습은 레이블이 없는 데이터를 사용하여 모델을 훈련시키는 방식입니다.

이 경우, 모델은 데이터 내 숨겨진 구조나 패턴을 스스로 찾아내야 합니다.



# 과적합 (Overfitting)

과적합은 모델이 훈련 데이터에 너무 잘 맞춰져서 새로운 데이터에는 잘 작동하지 않는 상태를 말합니다.

학생이 시험문제만 잘 푸는 것과 비슷하며, 이는 모델이 일반화되지 못했음을 의미합니다.





# 하이퍼파라미터 (Hyperparameter)

하이퍼파라미터는 모델의 학습 과정을 제어하는 외부 설정값입니다.

이는 모델의 구조와 훈련 방식에 영향을 주며, 튜닝을 통해 모델의 성능을 개선할 수 있습니다.



“주차장 프로젝트에 필요한 데이터는?”

## 데이터 수집





# 데이터 수집

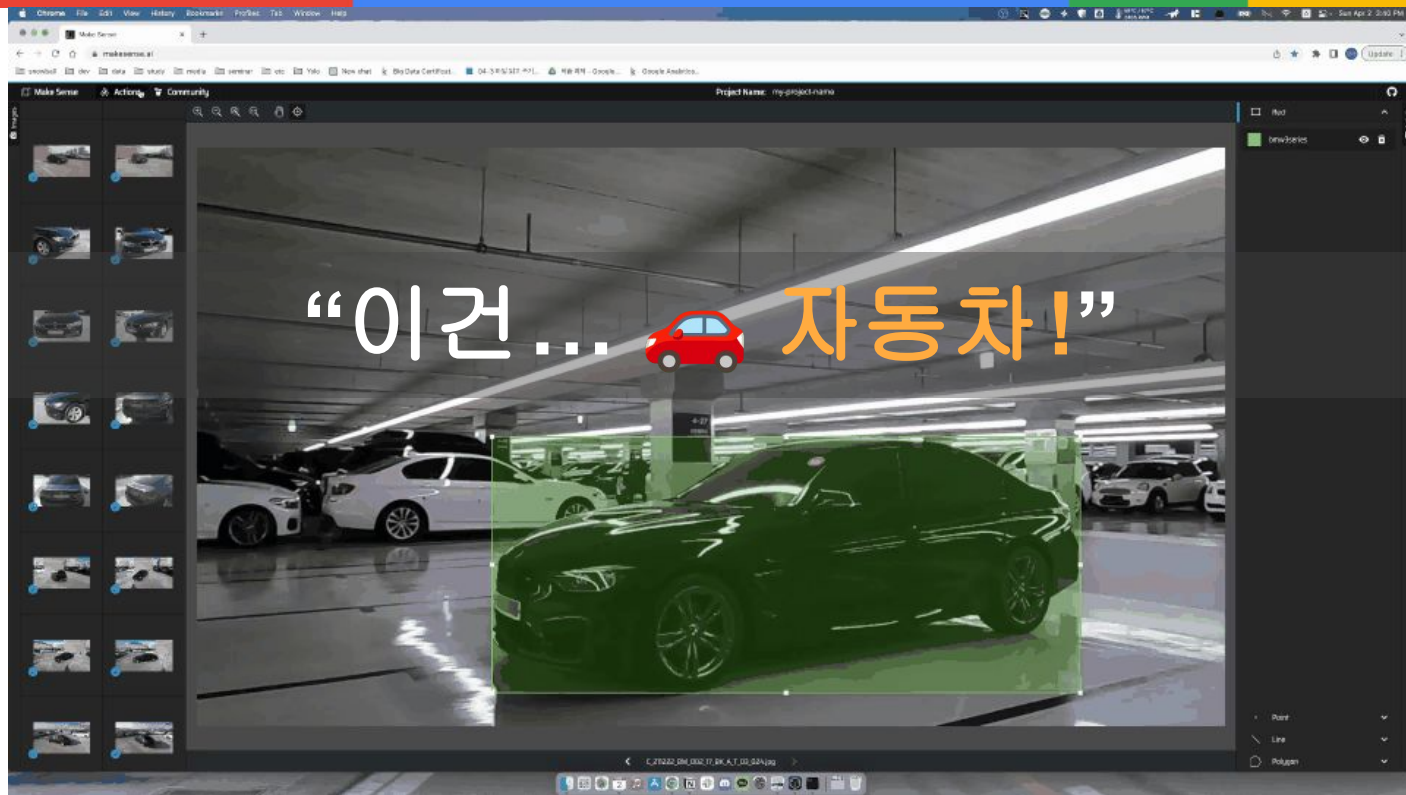




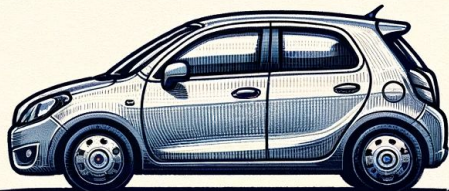
# 데이터 전처리



# 데이터 라벨링

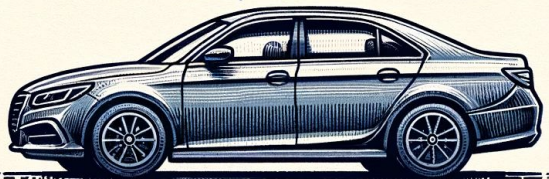


# 데이터 전처리



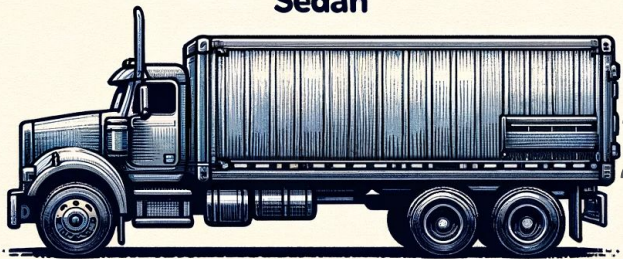
Compact Car

40장



Sedan

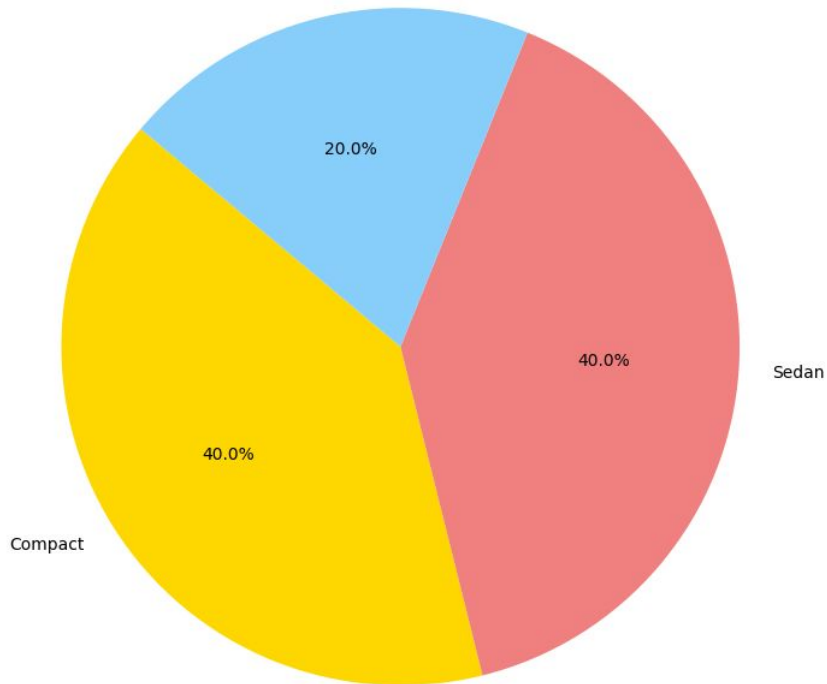
40장



Truck

20장

Vehicle Images Distribution  
Truck



# 데이터 전처리



**car\_dataset**

01\_car.png

02\_car.png

03\_car.png

04\_car.png

05\_car.png

...

100\_car.png



# 데이터 분리

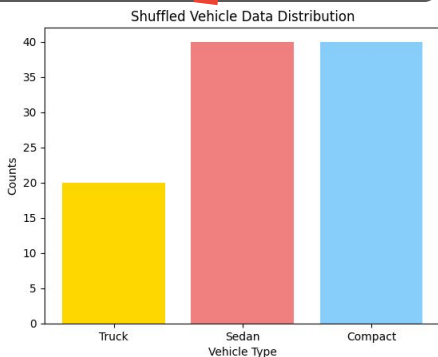
100장의 자동차 사진

80장

20장

학습 데이터

테스트 데이터



# 모델 과적합



“1 - 3 - 5 - 7 - 9 - ?? - ??”

학습 데이터

테스트 데이터

# 모델 과적합



“1 - 3 - 5 - 7 - 9 - 11 - 13 - ?? - ??”

학습 데이터

테스트 데이터

실제 시험

# 모델 훈련과 평가

## 1. 모델 훈련

- a. 마치 요리를 배우는 과정과 비슷합니다.
- b. 요리사(모델)가 다양한 재료(데이터)를 사용하여 요리(예측)하는 방법을 배웁니다.

## 2. 모델 평가

- a. 요리사가 만든 요리를 시식하여 맛(성능)을 평가하는 과정과 비슷합니다.
- b. 요리의 맛을 여러 기준(성능 지표)으로 평가합니다.





# 성능지표

1. 정확도(Accuracy)
2. 정밀도(Precision)
3. 재현율(Recall)



# 모델 평가 지표

## 혼동 행렬 (Confusion Matrix)

혼동 행렬은 모델의 예측값과 실제값을 비교하여 다음과 같은 네 가지 결과를 도출하는 표입니다:

- **True Positive (TP)**: 실제 Positive이고 예측도 Positive인 경우
- **False Positive (FP)**: 실제 Negative인데 예측은 Positive인 경우
- **False Negative (FN)**: 실제 Positive인데 예측은 Negative인 경우
- **True Negative (TN)**: 실제 Negative이고 예측도 Negative인 경우

### 실제값

### 예측값

	Positive	Negative
Positive	True Positive	False Positive
Negative	False Negative	True Negative

# 모델 평가 지표



## 정밀도 (Precision)

정밀도는 모델이 **Positive**로 예측한 것 중에서 실제로 **True**인 비율을 나타냅니다.





공식:  $\text{Precision} = \text{TP} / \text{TP} + \text{FP}$

## 재현율 (Recall)

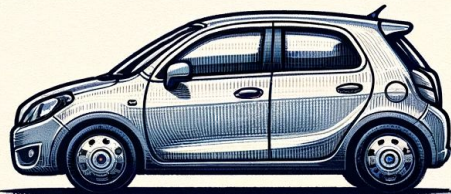
재현율은 실제로 **True**인 것 중에서 모델이 **Positive**로 예측한 비율을 나타냅니다.

공식:  $\text{Recall} = \text{TP} / \text{TP} + \text{FN}$

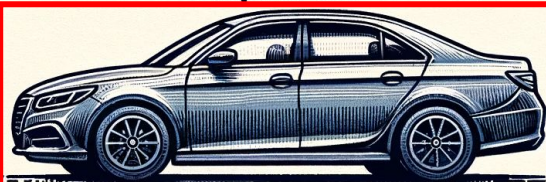
# 모델 평가

		Actual Values	
		1	0
Predicted Values	1	<b>TRUE POSITIVE</b> 	<b>FALSE POSITIVE</b>  <b>TYPE 1 ERROR</b>
	0	<b>FALSE NEGATIVE</b>  <b>TYPE 2 ERROR</b>	<b>TRUE NEGATIVE</b> 

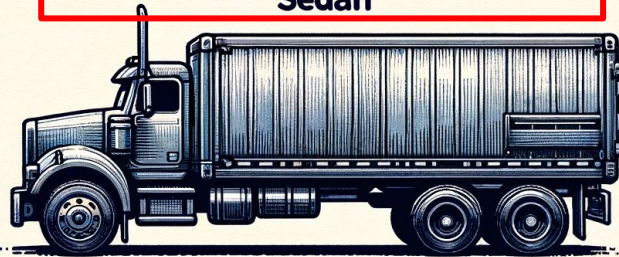
# 모델 평가



**Compact Car**



**Sedan**



**Truck**

# 모델 평가



실제값

예측값

	경차	승용차	트럭
경차	100%	0	0
승용차	0	100%	0
트럭	0	0	100%



# 모델 평가



실제값

예측값

	경차	승용차	트럭
경차	100%	0	0
승용차	100%	0%	0
트럭	0	50%	50%

# 모델 작업 절차

