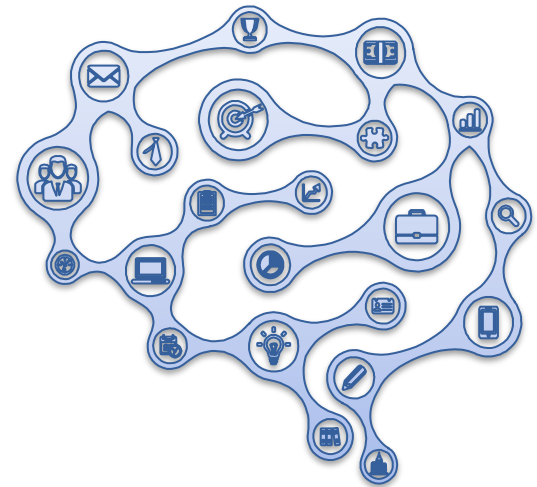


# 의료 Artificial Intelligence

지도 학습 / 비지도 학습 (chap7,8)

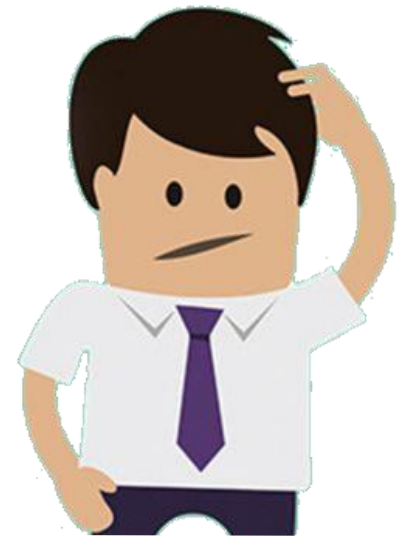
2022.05.26



# 오늘 배울 내용 ...

1. 지도학습과 비지도학습
2. 인공지능 실습
3. mblock 실습

어렵지 않다  
쉬운 것도 아니다



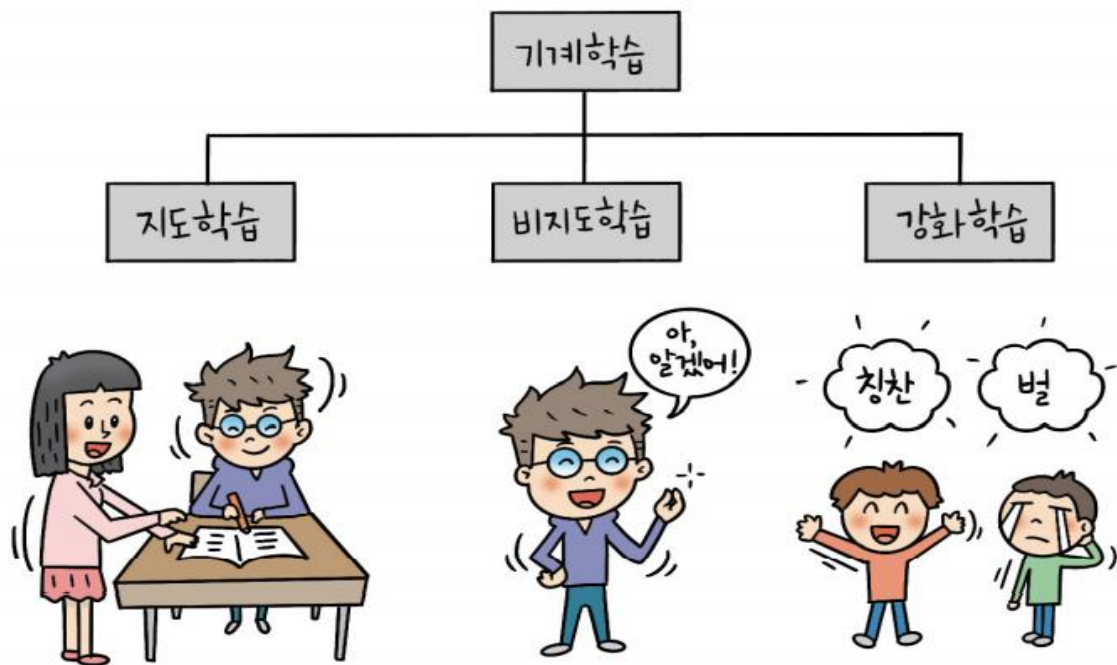
---

# 인공지능 이론

---

# 머신러닝과 구분

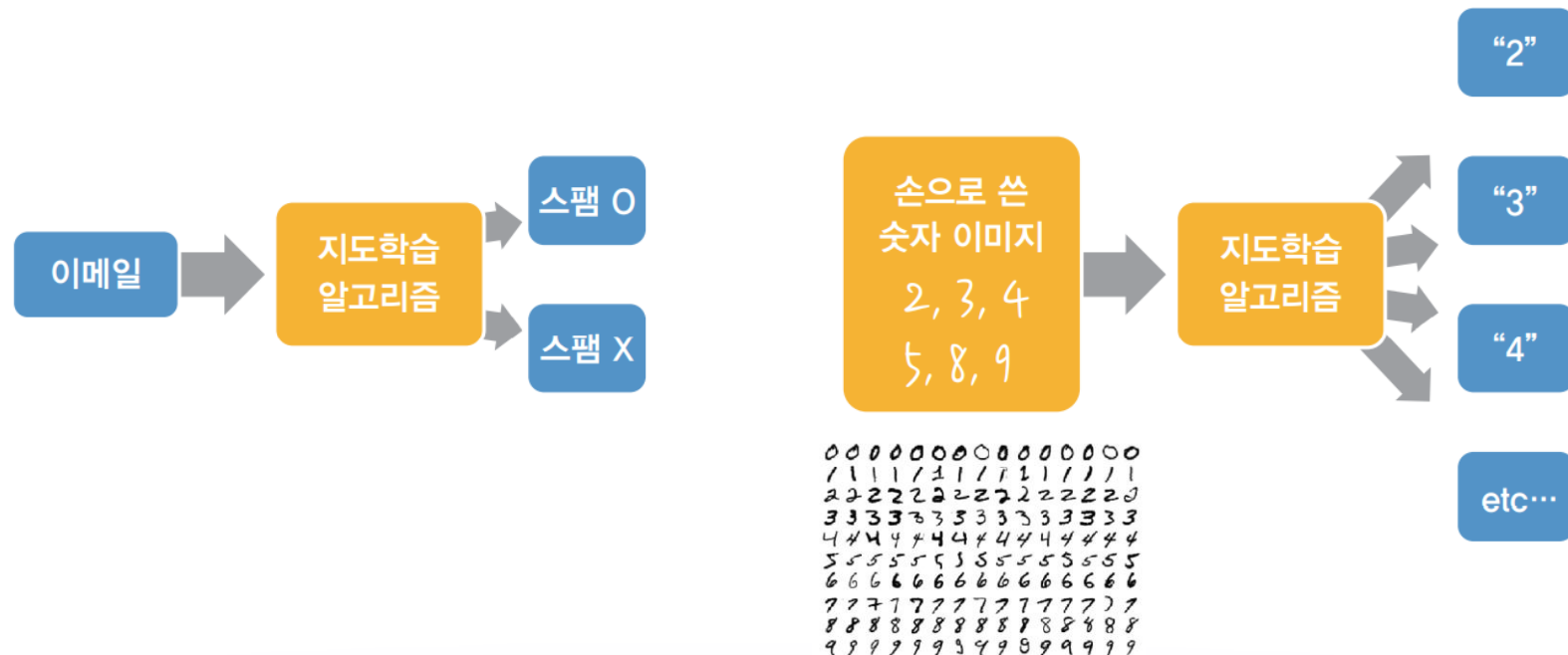
- 입력하는 데이터에 레이블(명시적인 정답)을 주느냐, 아니냐에 따라 지도학습, 비지도학습으로 구분
- 강화학습은 행동(상태)의 변화 후에 환경으로부터 보상을 받아 학습



▲ 기계학습의 유형

# 지도학습(Supervised Learning)

- 문제(Feature)와 정답(Label, Annotation)이 있는 데이터(Training Set)를 컴퓨터에 학습시킨 후 새로 입력된 데이터(Test Set)를 분류하거나 예측하는 학습 방법
- 입력 데이터  $x$ 와 그에 대한 정답 레이블 Label  $y$ 의 쌍 Pair  $(x, y)$ 를 이용해서 학습
- 지도학습의 목적은 입력 데이터와 결괏값을 이용해 특정한 타깃을 예측하는 것



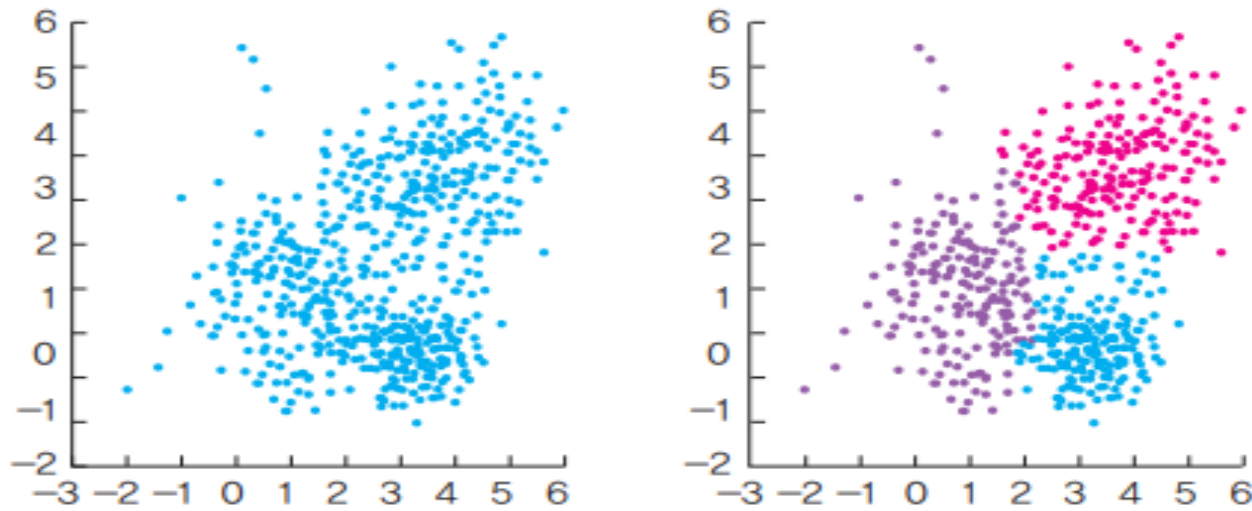
# 비지도 학습(Unsupervised Learning)

- 지도학습과 달리 훈련 데이터에 정답 레이블이 없음.

입력된 데이터를 받아 컴퓨터가 스스로 데이터의 특징을 찾아 내는 것

- 비지도학습은 데이터의 숨겨진 특징(Hidden Feature)을 찾아내는 것에 목적

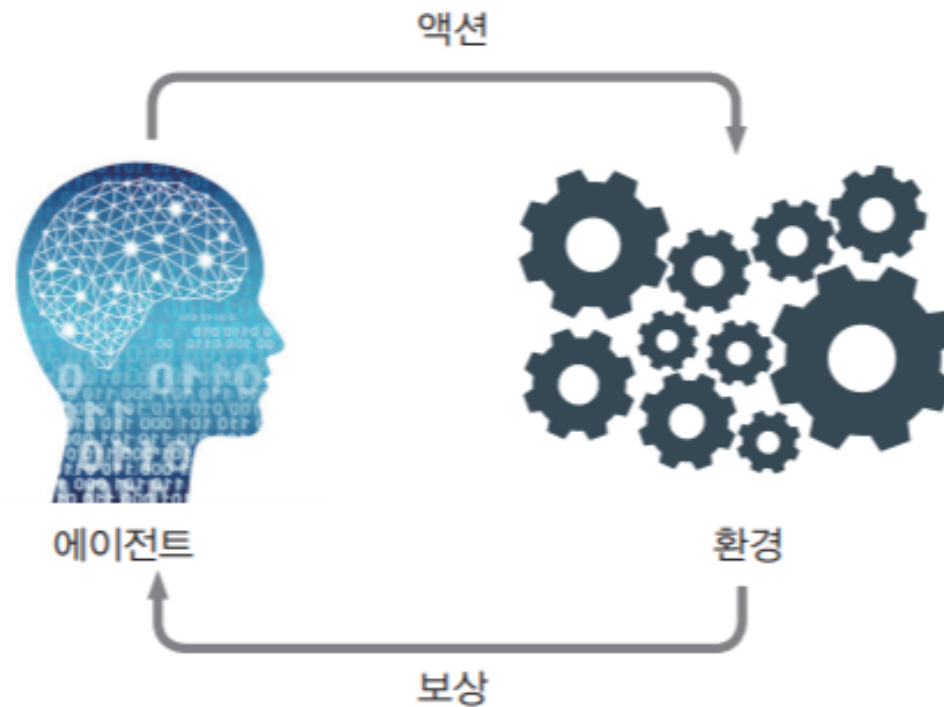
. 비지도학습의 목표는 데이터의 특성을 분석해서 유사한 속성을 가진 값끼리 그룹화(군집, Clustering)  
하고 그룹의 특징을 찾아 냄



▲ 가공되지 않은 데이터와 클러스터화된 데이터

# 강화학습

- 강화학습은 준지도학습으로 행동한 결과(처리 결과)에 대해 보상(평가) 알고리즘을 제공하여 스스로의 행동을 개선하는 방법



# 지도학습 알고리즘

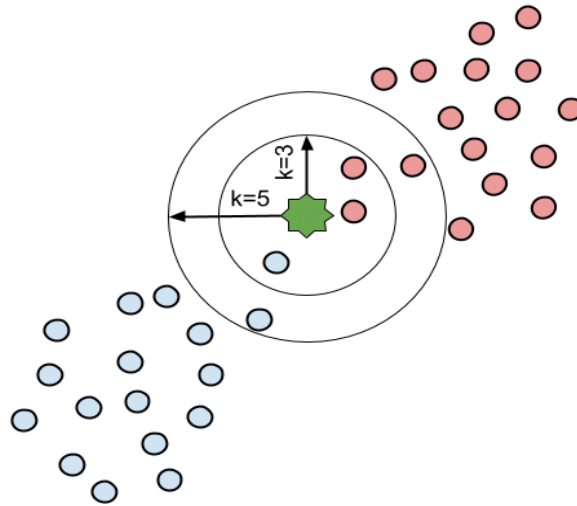
유형	<p><b>k-최근접 이웃(k-Nearest Neighbors)</b> 새로운 데이터가 무엇인지 구분하기 위해 훈련에 사용됐던 데이터에서 가장 가까운 데이터의 유형, 즉 최근접 이웃을 찾음</p>
	<p><b>로지스틱 회귀(Logistic Regression)</b> 선형 함수 결과를 시그모이드 함수를 이용해 0~1 사이로 압축한다. 로지스틱 회귀는 다중 분류도 지원한다. - Binary Classification(이진 분류), Multiclass Classification(다중 분류) - 스팸 분류, 암 진단, 꽃의 품종 판별, 손글씨 숫자 분류</p>
	<p><b>서포트 벡터 머신(Support Vector Machine)</b> 두 카테고리 중 어느 하나에 속한 데이터의 집합이 주어졌을 때, 새로운 데이터가 어느 카테고리에 속할지 판단하기 위해 데이터들의 경계 중 가장 큰 폭을 가진 경계를 찾는 알고리즘</p>
	<p><b>나이브 베이즈(Naive-Bayes)</b> 복잡한 베이즈 정리를 간단한 조건부 확률 방법을 이용해 분류하는 알고리즘. 나이브 베이즈 분류는 텍스트 분류에 유용하게 사용되며 스팸메일 필터, 텍스트 분류, 감정 분석, 추천 시스템에서 효과적이다.</p>
	<p><b>의사결정 트리(Decision Trees)</b> 스무고개 하듯 예/아니요 질문을 이어가며 특정 기준(질문)에 따라 데이터를 구분하는 학습 모델. 신용 등 급 분류, 식물 분류, 환자 판별 등에 사용.</p>
	<p><b>랜덤 포레스트(Random Forest)</b> 분류, 회귀분석 등에 사용되는 앙상블 학습 방법의 일종으로, 훈련 과정에서 구성한 다수의 결정 트리로부터 분류 또는 평균 예측값(회귀분석)를 출력함으로써 과적합을 막는 학습이 가능</p>
	<p><b>신경망/딥러닝(Neural Networks/Deep Learning)</b> 인간의 신경 모델을 모방한 학습 방법으로, 여러 개의 층으로 이뤄진 뉴런들의 가중값을 구함. 딥러닝의 등장으로 이미지 인식, 언어 처리 등 인공지능의 전성기를 이루고 있는 알고리즘</p>



# k-최근접 이웃 알고리즘

- k-NN: 새로운 데이터가 주어질 때 기존 데이터 가운데 가장 가까운 k개 이웃의 정보로 새로운 데이터 예측

- 학습 절차 없음

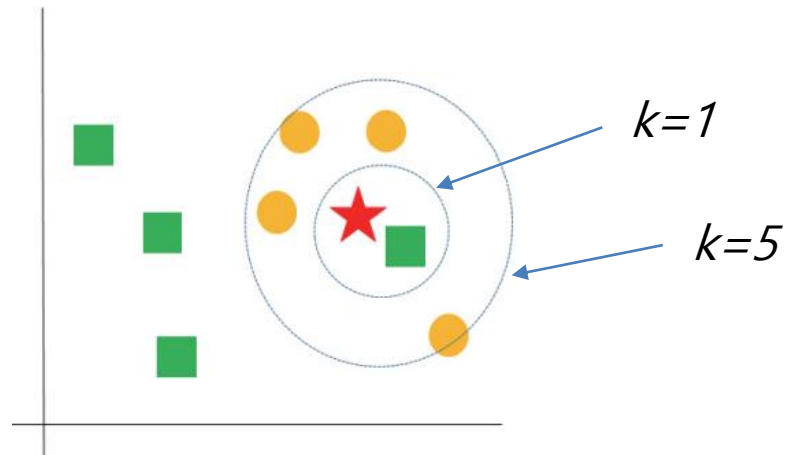


# k-최근접 이웃 알고리즘

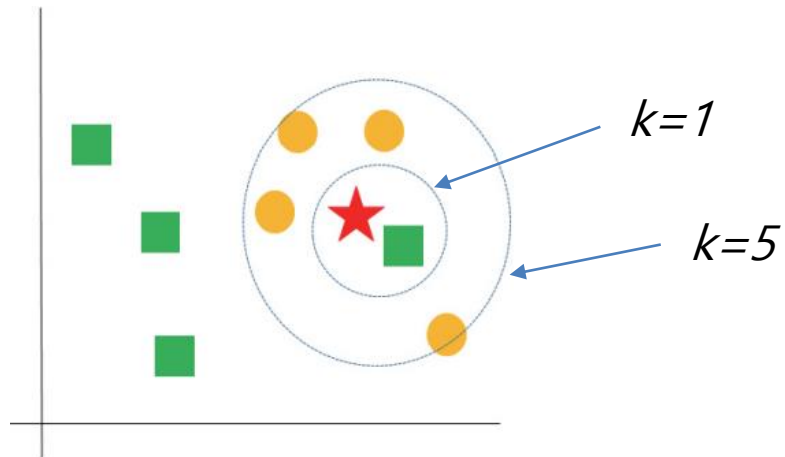
- 가장 가까이에 있는 단일 자료만 보는 것이 아니라 주변에 분포돼 있는 데이터들의 거리를 측정해 가장 많이 이웃한 것으로 분류하는 방법을 사용
- $k$ 는 분류를 판단하기 위해 참조할 주변의 데이터 수를 의미함.

작은 원은  $k=1$  일 경우 → 가장 가까운 1개로 분류를 판단하겠다는 것

두 번째 원은  $k=5$  일 경우 → 가장 가까운 5개의 데이터로 분류를 판단하겠다는 것



# k-최근접 이웃 알고리즘



· k가 1일 경우★은■에 속하는 것으로 분류

· k가 5일 경우★은●에 속하는 것으로 분류

· k의 값이 커질수록 분류 자체를 하지 못할 수 있으므로 가능한 한 작은 수 사용

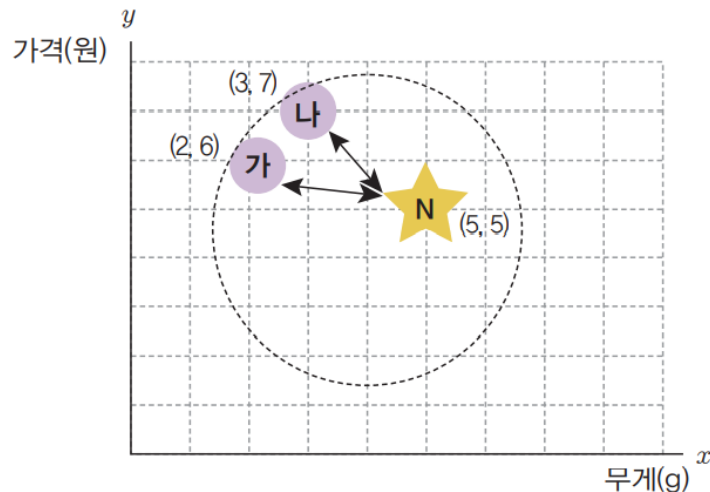
· k값은 기본적으로 홀수를 사용

# k-최근접 이웃 알고리즘

- k-NN은 거리 측정 방법에 따라 결과가 달라지므로 표준화된 거리 측정 방법이 중요.  
주로 유클리드 거리로 계산

$$\sqrt{(Ax - Bx)^2 + (Ay - By)^2}$$

- 상품 판매에서 무게와 가격으로 배치되었을 때 다음 그림의 y좌표 단위가 달러일 경우 새로운 상품의 유사도를 구하기 위한 유클리드 거리



가와 N의 유클리드 거리는  $\sqrt{(5-2)^2 + (5-6)^2} = \sqrt{9+1} = \sqrt{10} = 3.162$

나와 N의 유클리드 거리는  $\sqrt{(5-3)^2 + (5-7)^2} = \sqrt{4+4} = \sqrt{8} = 2.828$

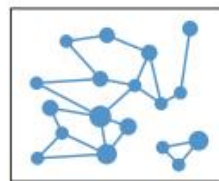
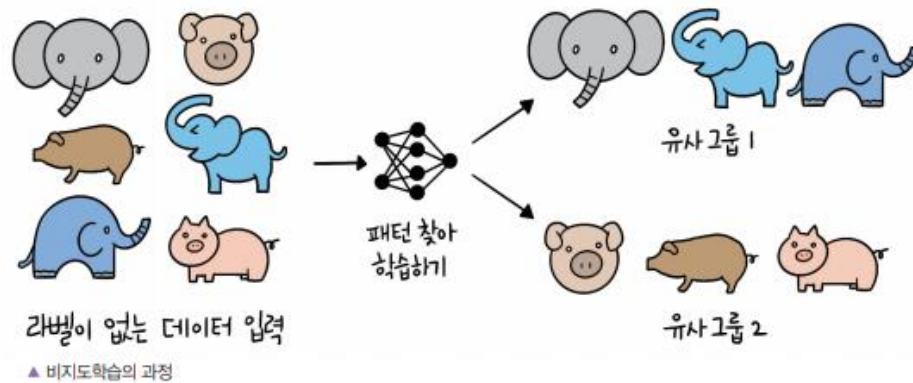
그러므로 나와 N이 가깝다.

# k-최근접 이웃 알고리즘

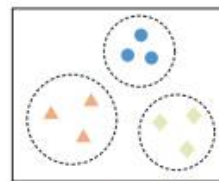
- 장점: 원리가 간단해 구현하기 쉬움.
  - 수치 기반 데이터에서 분류, 비슷한 것 찾기 등에서 많이 활용
- 단점 : 분석할 때 적합한 최적 이웃의 수와 거리 척도의 방법을 연구자 임의로 정해야 하고  
새로운 관측 값과 학습 데이터 사이의 거리를 전부 측정해야 하므로 시간이 오래 걸림
- 활용 : 얼굴 인식, 개인 영화 추천, 단백질 및 질병 추출을 위한 유전자 데이터 패턴 식별 등에 활용

# 비지도 학습

- 비지도학습은 알고리즘에 입력 데이터만 제공하고 데이터의 패턴을 스스로 식별.
- 목표는 알고리즘이 데이터에서 기본 패턴 또는 특징을 식별해 이를 더 잘 이해하는 것
- 지도학습의 입력 데이터(데이터 - 군집결과)로 활용해서 지도학습의 성능을 끌어올리는 용도로 활용



이상 징후 감지



군집

▲ 비지도학습의 유형

# 비지도학습의 유형

군집	<b><u>k-평균(k-means)</u></b> : 주어진 데이터를 지정된 클러스터 갯수(k)로 그룹핑. 한 클러스터 내의 데이터들은 동일한 성질을 가지며 다른 그룹에 대해 구별됨.
	<b>계층 군집 분석(Hierarchical Cluster Analysis, HCA)</b> : 계층 군집은 유사한 군집끼리 묶어가면서 최종 적으로는 하나의 케이스가 될 때까지 군집을 묶는 클러스터링. 군집 간의 거리를 기반으로 클러스터링을 하는 알고리즘이며 K-평균과는 다르게 군집의 수를 미리 정해주지 않아도 됨.

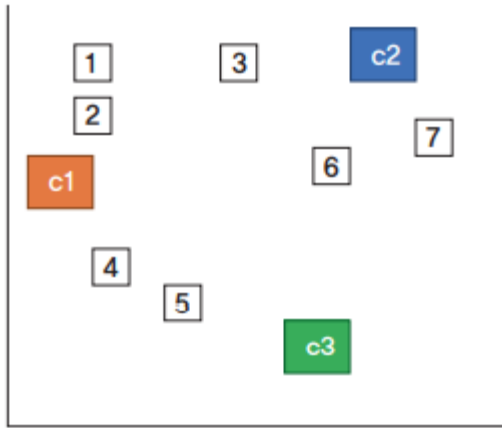
# k-평균 알고리즘 주요 용어

- 클러스터: 유사한 특성(중심을 기준으로 가까운 것)을 가진 데이터끼리의 묶음
- 클러스터링: 어떤 데이터들이 주어질 때, 그 데이터들을 클러스터로 무리 지어 주는 것
- 센트로이드(Centroid): 무리, 즉 클러스터의 중심
- k: 몇 개의 무리로 클러스터링할 것인지. 데이터를 3개로 클러스터링한다면  $k = 3$

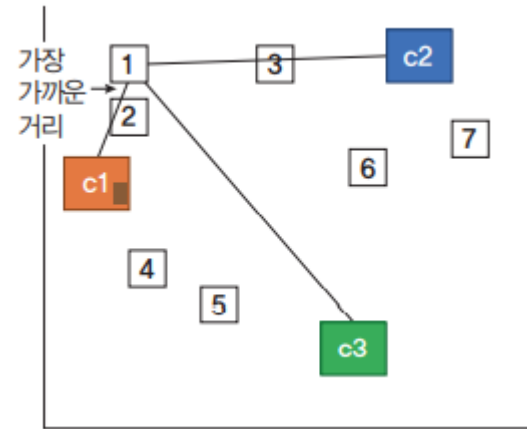


# k-평균 알고리즘 과정

7개의 데이터가 있고,  $k=3$ 으로 가정



① 무작위로 3개의 센트로이드를 설정



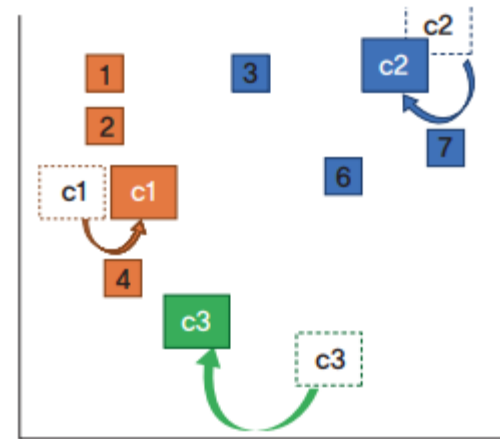
② 데이터 하나씩 가장 가까이에 있는 센트로이드를 찾는다.

1의 데이터의 경우 3개의 센트로이드 중 c1이 가장 가까우므로 c1으로 클러스터링한다. 이런 방법으로 데이터와 센트로이드 간의 거리를 측정해 클러스터링한다.

# k-평균 알고리즘 과정

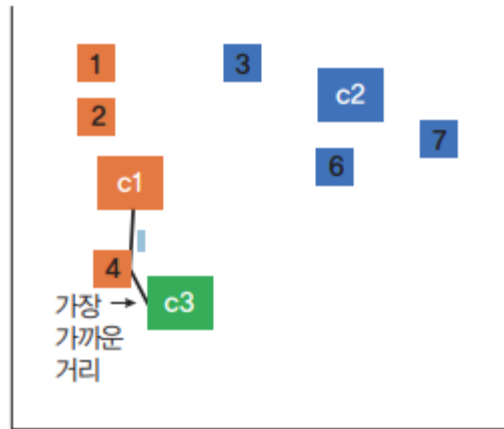


③ 앞의 2번 과정의 결과 위 그림처럼 클러스터링 됨

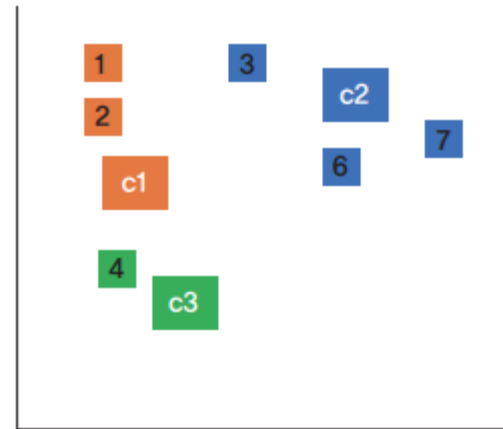


④ 각 센트로이드가 클러스터링 된 데이터 간의 평균거리로 중심점을 옮긴다

# k-평균 알고리즘 과정

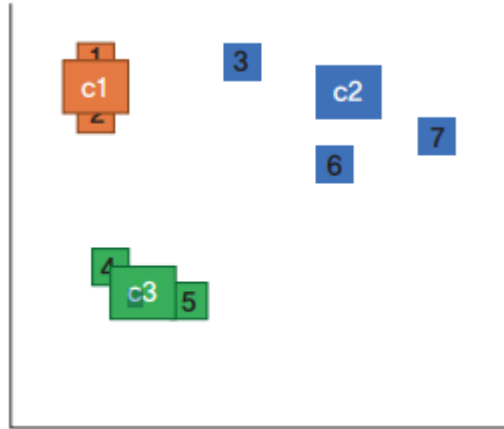


⑤ 센트로이드의 중심점이 옮겨졌으므로 각 데이터의 센트로이드와의 거리를 다시 측정

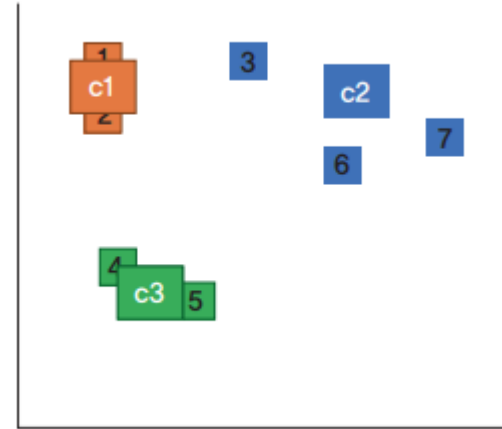


⑥ 데이터 4의 경우 c3과 가까워져 클러스터가 바뀌고 나머지 데이터들은 변경이 없다.

# k-평균 알고리즘 과정



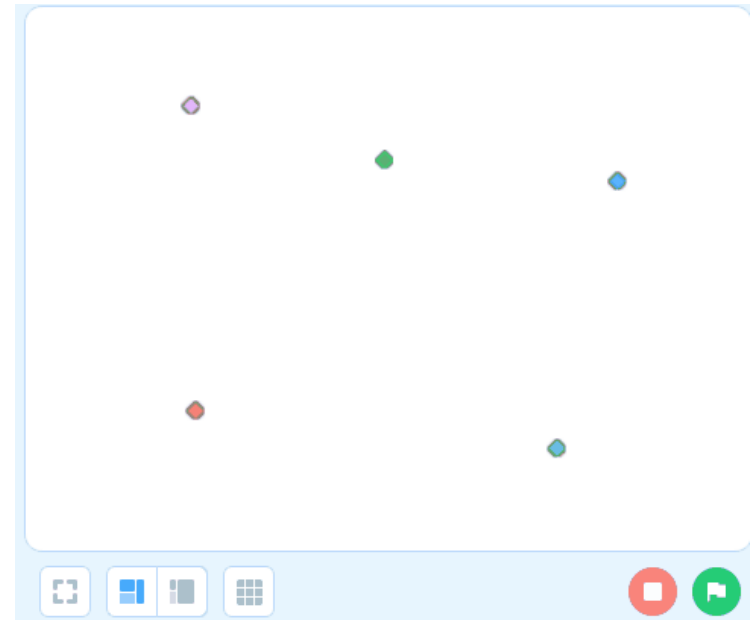
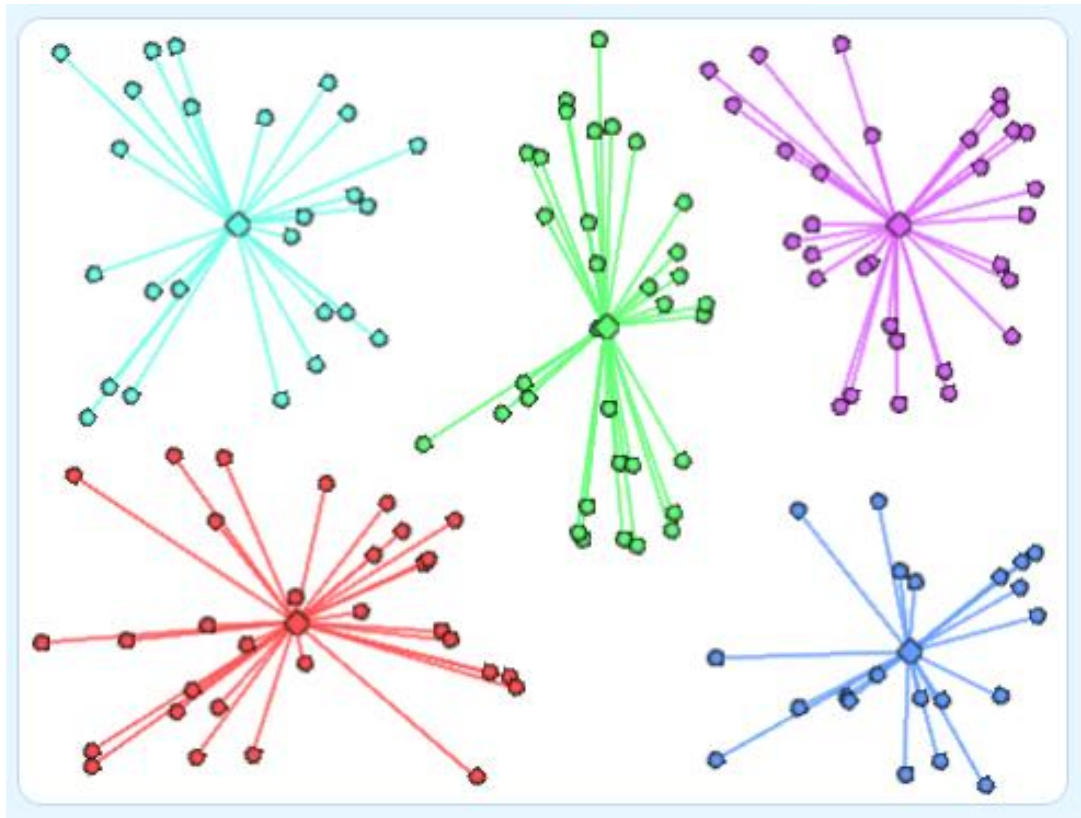
⑦ 다시 한번 센트로이드를 중심으로  
각 데이터 간의 평균 거리를 구하고  
중심점을 옮긴다.



⑧ 이러한 방법으로 데이터와 센트로이드 간  
의 거리를 구하고 중심점이 더 바뀌지 않을  
때까지 반복한다.

# k-평균 알고리즘 과정

mblock에서 day12\_ex1프로젝트 실행




# 인공지능 실습

# 인공지능 실습

<https://animalface.site/ko/index.html>

성별을 선택하세요

여자 ☒ 남자



**포근한 매력의 곰상**

첫 인상은 무서워 보이지만 알고 보면 귀여운 매력의 당신! 꼼꼼하고 섬세한 성격으로 연인을 헌신적으로 챙겨주는 당신은 연인에게 듬직한 존재! 포근한 매력에 듬직할까지 갖춘 최고의 남자대!

곰상 연예인: 마동석, 조진웅, 조세호, 안재홍

곰상	99%
곰룩상	1%
강아지상	0%
토피상	0%
고양이상	0%

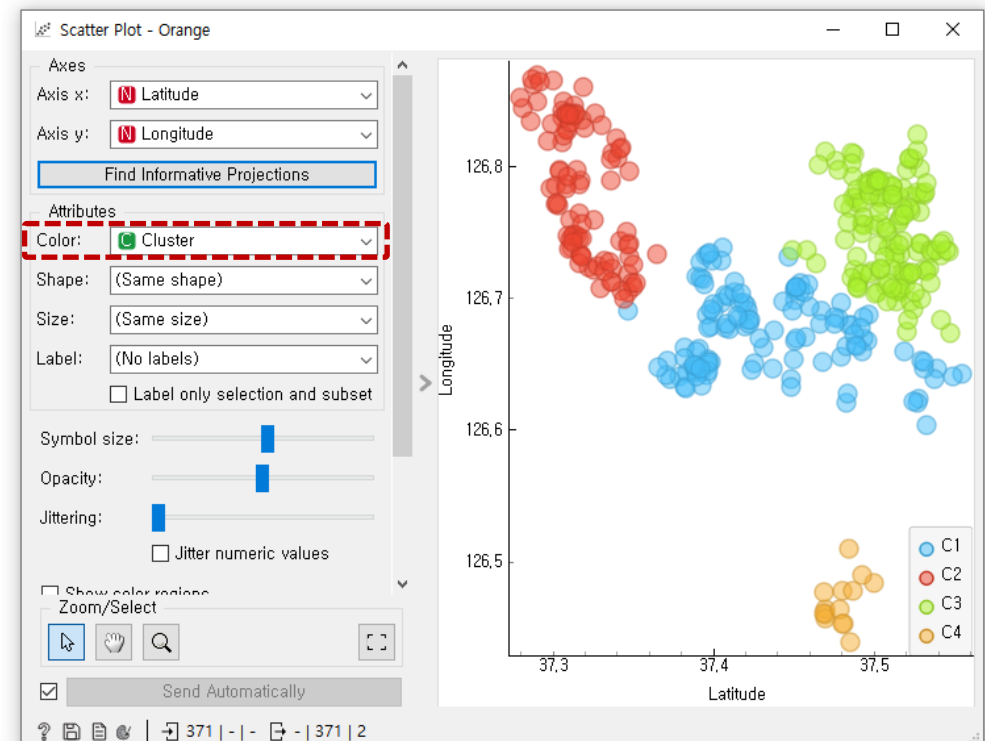
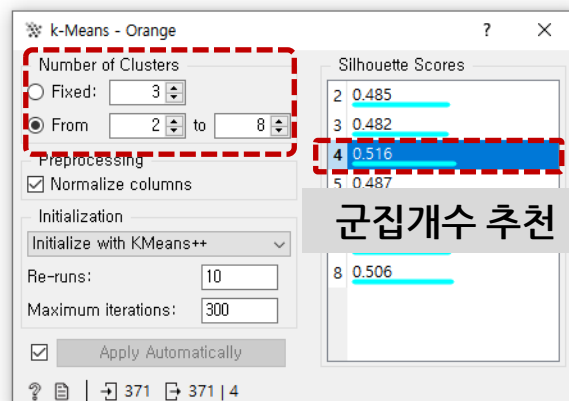
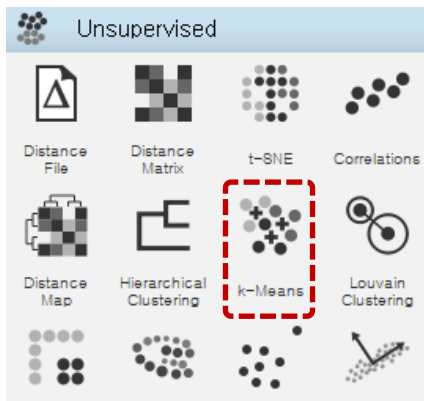
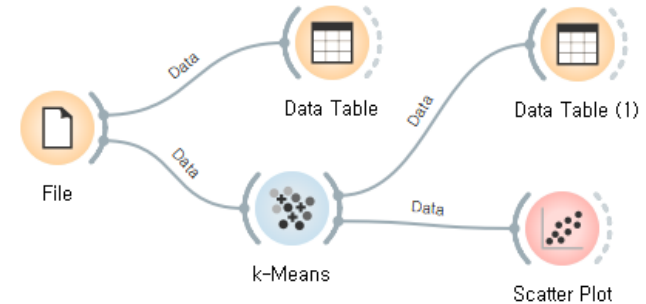
🔔 🌐 🏠 + 63.2K

다른 사진으로 재시도

# 인공지능 실습

## 실습 1

- 데이터 불러오기 : 'Delivery.csv'
- 데이터 모델 : k-Means

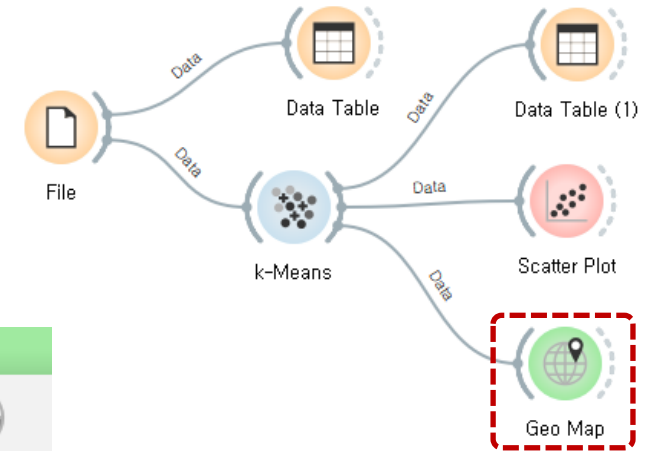
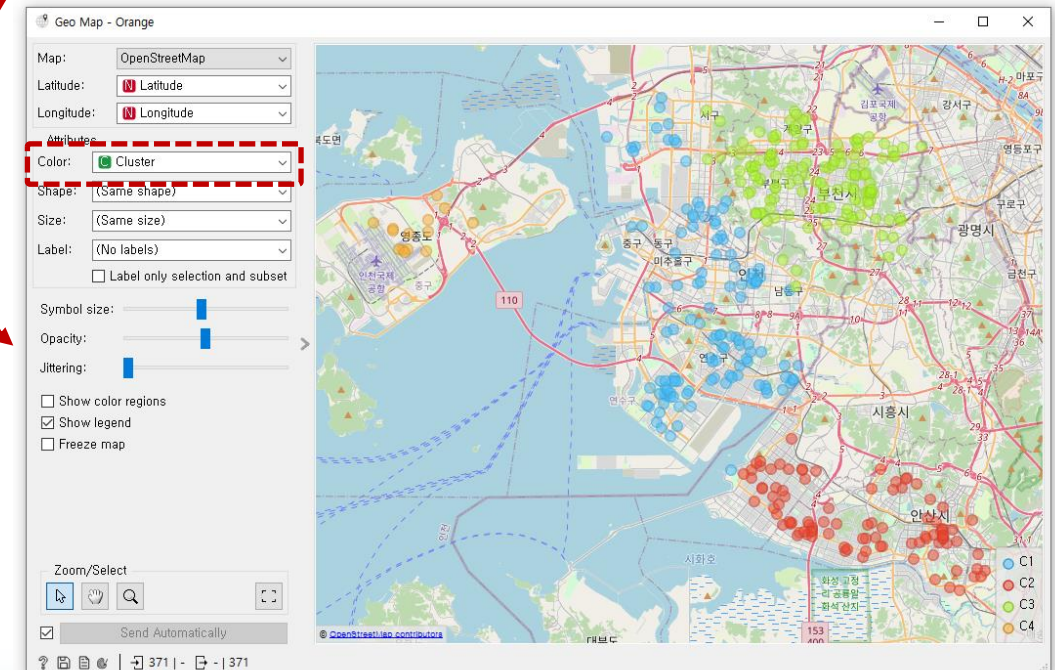
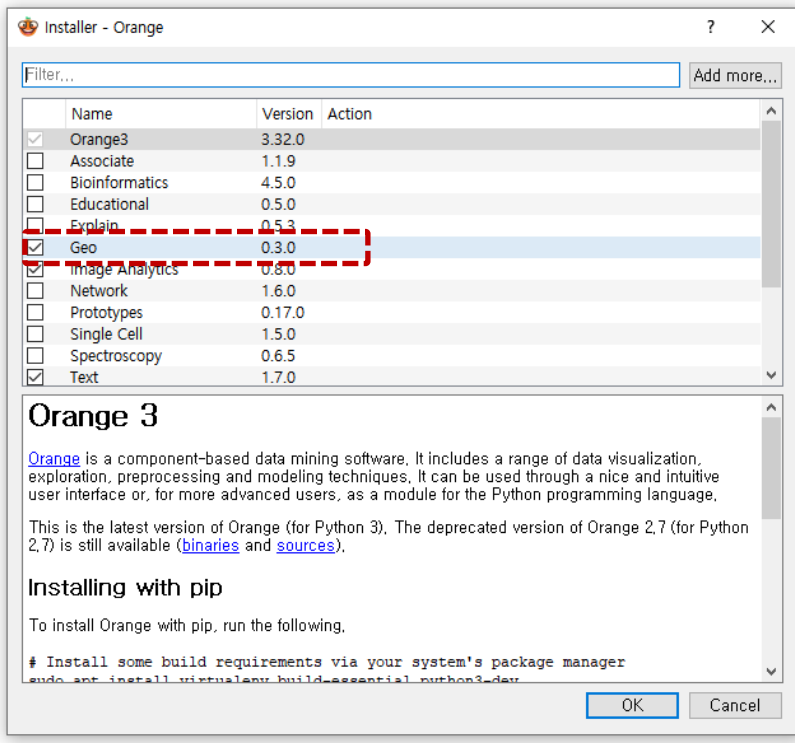
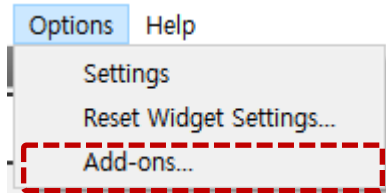




# 인공지능 실습

## 실습 1-2 : 지도상에 표시

- 데이터 불러오기 : 'Delivery.csv'
- 데이터 모델 : Geo-



---

# MBlock 실습

---

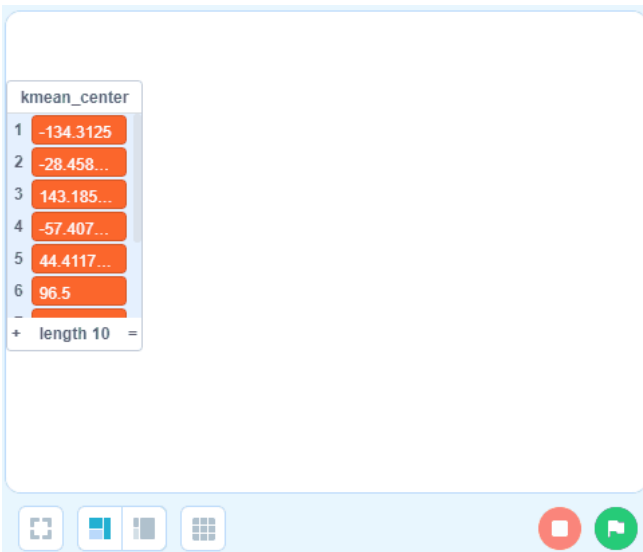
# MBlock 인공지능 실습 1

## day12\_ex1 프로젝트 불러 오기



kmean\_center

클러스터...



클러스터...

클러스터...

클러스터...

클러스터...

이동 ▾ 을(를) 받았을 때  
0.5 초 동안 x: cluster\_place ▾ 의 1 번째 항목 y: cluster\_place ▾ 의 2 번째 항목 로(으로) 이동하기  
kmean\_center ▾ 에 1 번째 항목을 cluster\_place ▾ 의 1 번째 항목 로 바꾸기  
kmean\_center ▾ 에 2 번째 항목을 cluster\_place ▾ 의 2 번째 항목 로 바꾸기

이동 ▾ 을(를) 받았을 때  
0.5 초 동안 x: cluster\_place ▾ 의 3 번째 항목 y: cluster\_place ▾ 의 4 번째 항목 로(으로) 이동하기  
kmean\_center ▾ 에 3 번째 항목을 cluster\_place ▾ 의 3 번째 항목 로 바꾸기  
kmean\_center ▾ 에 4 번째 항목을 cluster\_place ▾ 의 4 번째 항목 로 바꾸기

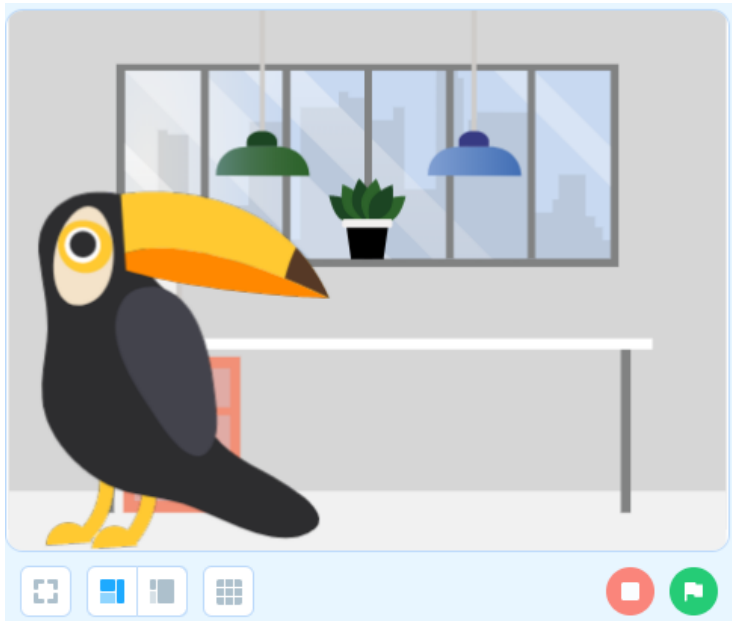
이동 ▾ 을(를) 받았을 때  
0.5 초 동안 x: cluster\_place ▾ 의 5 번째 항목 y: cluster\_place ▾ 의 6 번째 항목 로(으로) 이동하기  
kmean\_center ▾ 에 5 번째 항목을 cluster\_place ▾ 의 5 번째 항목 로 바꾸기  
kmean\_center ▾ 에 6 번째 항목을 cluster\_place ▾ 의 6 번째 항목 로 바꾸기

이동 ▾ 을(를) 받았을 때  
0.5 초 동안 x: cluster\_place ▾ 의 7 번째 항목 y: cluster\_place ▾ 의 8 번째 항목 로(으로) 이동하기  
kmean\_center ▾ 에 7 번째 항목을 cluster\_place ▾ 의 7 번째 항목 로 바꾸기  
kmean\_center ▾ 에 8 번째 항목을 cluster\_place ▾ 의 8 번째 항목 로 바꾸기

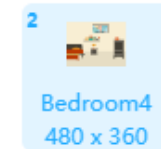
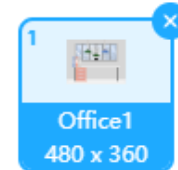
이동 ▾ 을(를) 받았을 때  
0.5 초 동안 x: cluster\_place ▾ 의 9 번째 항목 y: cluster\_place ▾ 의 10 번째 항목 로(으로) 이동하기  
kmean\_center ▾ 에 9 번째 항목을 cluster\_place ▾ 의 9 번째 항목 로 바꾸기  
kmean\_center ▾ 에 10 번째 항목을 cluster\_place ▾ 의 10 번째 항목 로 바꾸기

# MBlock 인공지능 실습 2

## 얼굴인식 출입기



코스튬 수: 2



# MBlock 인공지능 실습 2

모양

소리

+

확장

검색하기

스프라이트 확장    디바이스 확장

AI   AI   Data Science   IOT   ART

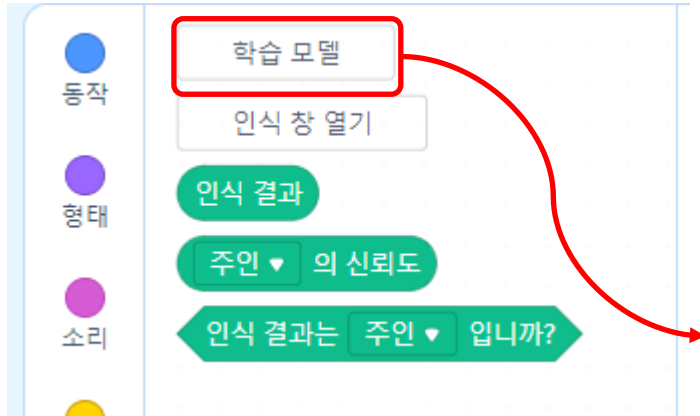
 <b>인식 서비스</b> By mBlock official 인식 서비스 API를 사용하면 비디오, 음성, 언어 및 지식과 같은 다른 기능들을 추가할 수 있습니다. <a href="#">더 보기</a> X 지음	 <b>기계학습</b> By mBlock official 기계 학습을 사용하면 프로그래밍할 필요는 없지만 컴퓨터가 학습하여 일을 배우고 인간의 두뇌와 유사한 인공 신경 네트워크를 구축 X 지음	 <b>데이터 차트</b> By mBlock official 이 확장을 사용하면 여 데이터를 시각화할 수 있습니다. 자르는 긴 단어보다 더 크게 알린다. + 추가	 <b>사용자 클라우드 메시지</b> By mBlock official 클라우드 메시지 확장을 사용하면, 계정의 데이터를 다양한 기기와 프로젝트로 동기화할 수 있습니다. + 추가	 <b>펜</b> 개발자: MIT Media Lab 당신의 스프라이트를 그립니다. X 지음	 <b>음악</b> 개발자: MIT Media Lab 악기와 드럼을 연주합니다. + 추가	 <b>기후 데이터</b> By mBlock official 확장은 실시간 기후 데이터에 대한 빠른 액세스를 제공합니다. + 추가	 <b>업로드 모드 브로드캐스트</b> By mBlock official 이 확장을 추가하여 확장자가 업로드 모드에서 스프라이트와 상호 작용할 수 있도록 합니다. + 추가
 <b>Google 스프레드시트</b> By mBlock official 이 확장을 사용하면 mBlock을 사용하여 Google 시트에 데이터를 입력할 수 있습니다. (Google 서비스 지역에서만 사용 가능) + 추가	 <b>mBlock 3 blocks</b> By mBlock official A collection of coding blocks which are included in mBlock 3 while not owned by mBlock 5. + 추가	 <b>小学教师工具箱</b> 개발자: Juangm2005 小学教师实用工具包。分学科，不断更新。小学数学：基础知识的值。操作的值 小学语文：课文背诵（设定背诵，删除所有背诵，删除 + 추가	 <b>AI Service</b> By mBlock official The extension is only available in China. Use Baidu AI service to recognize images, texts, speeches, human body, and process natural + 추가	 <b>Video Sensing</b> 개발자: MIT Media Lab Sense motion with the camera. + 추가	 <b>Makey Makey</b> 개발자: MIT Media Lab Make anything into a key. + 추가	 <b>Text to Speech</b> 개발자: MIT Media Lab Make your projects talk. (not available in China yet) X 지음	 <b>Translation</b> By mBlock official Translate your texts into different languages. + 추가

TM

인식 서비스

텍스트 음성 변환

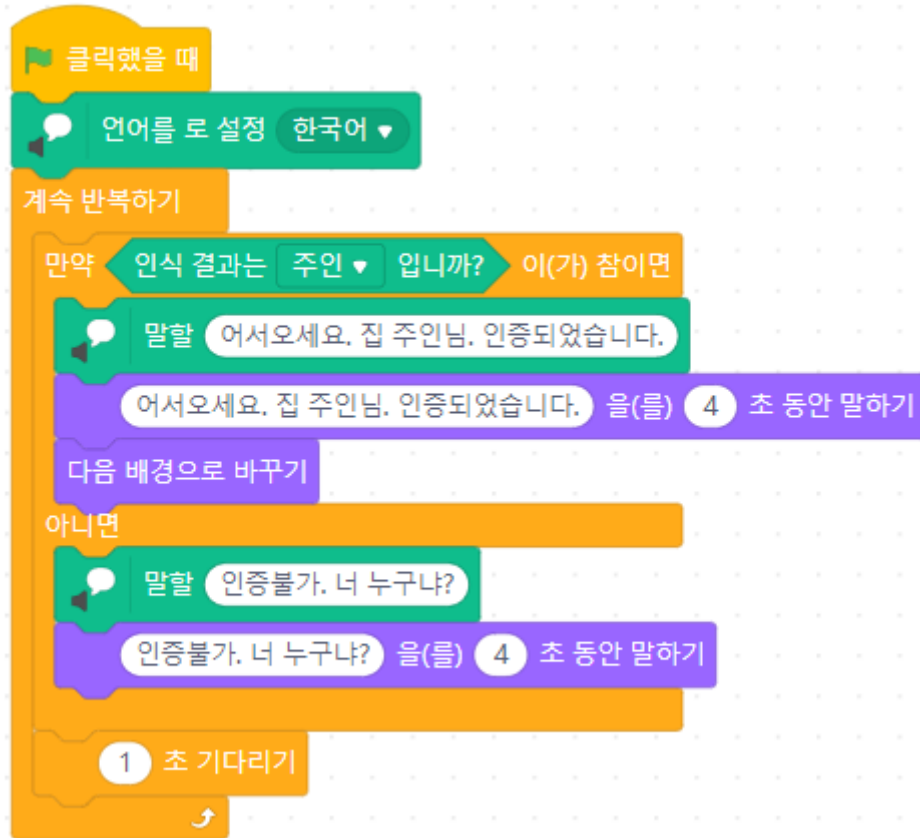
# MBlock 인공지능 실습 2



The '모델 학습' (Model Learning) window in MBlock v5.4.0 displays a live video feed from the 'HD WEB CAMERA (0a50:6100)'. It shows three examples of object recognition with their respective confidence levels and '배우기' (Learn) buttons. The first example is '42 예시' (Example 42) for '주인' (Owner) with a 90.0% confidence level. The second example is '23 예시' (Example 23) for '주인친구' (Owner's friend) with a 10.0% confidence level. The third example is '21 예시' (Example 21) for '주인친구2' (Owner's friend 2) with a 10.0% confidence level. At the bottom, there are buttons for '새로운 모델 만들기' (Create new model) and '모델 사용' (Use model), with the latter highlighted by a red box. A '결과 주인' (Result Owner) box is also visible on the right.

예시 (Example)	주인 (Owner)	신뢰도 (Confidence)	배우기 (Learn)
42 예시	주인	90.0%	배우기
23 예시	주인친구	10.0%	배우기
21 예시	주인친구2	10.0%	배우기

# MBlock 인공지능 실습 2



---

팀 활동

---