
인공지능 이론 문제

1. 통계학에서 '표본의 크기가 충분히 크다면, 표본평균들의 분포는 정규분포를 이룬다'는 정리는 무엇인가?.

2. 표준정규분포에서 Z값에서 확률변수(X)를 구하는 공식을 쓰시오.

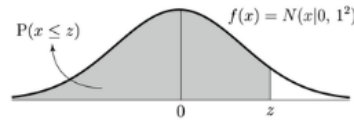
(μ : 평균, σ : 표준편차)

3. 평균 $\mu=150$ 이고, 표준편차 $\sigma=10$ 인 정규분포 $N(150, 10^2)$ 에 대해 $X=170$ 에 대한 Z 점수를 계산하시오.

→ 2

4. 한 회사의 건전지 수명은 평균이 110시간, 표준편차가 10인 정규분포를 따른다고 할 때 건전지 수명이 100시간 이상 120시간 이하일 확률은?

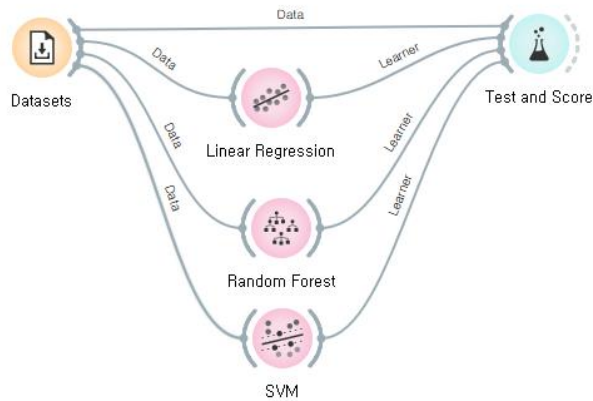
→ 0.6826



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857

5. 데이터마이닝에 대해서 정의하시오.
6. 데이터 마이닝 과정에서 요약이나 집계등을 수행하여 데이터마이닝 (분석)을 위한 적합한 형태로 데이터를 가공하는 과정을 무엇이라고 하는 가?
7. 데이터형태 중 반정형(semi-structured) 데이터 형식 중 대표적인 것 2가지를 제시하시오 ?
8. 데이터마이닝 모델 중 특정 변수에 대해 어떠한 숫자값을 예측하는 모델은?
9. 데이터마이닝 모델에서 군집화는 유사도에 따라 비슷한 개체를 하나의 군집으로 묶는 데 유사도는 무엇으로 측정하는 가?
10. 데이터마이닝 중 회귀분석은 어떤 마이닝 방법인지 기술하시오
11. 선형 회귀분석에서 요인이 되는 변수에 대해 결과값으로 나타나는 결과 데이터를 무슨 변수라고 하는 가?
12. 머신러닝에 대해 정의하시오
13. 머신러닝을 수학적으로 정의하면 (), (), 군집 이라고 할 수 있다. ()에 들어갈 용어는?
14. '머신러닝 프로그램은 어떤 ()에 관련된 ()로 부터 ()를 가지고 학습하는 컴퓨터 프로그램을 말한다'
상기 정의에서 3군데 괄호안에 들어가는 용어를 쓰시오.

15. 데이터마이닝을 통해서 다음과 같이 모델을 만들었다. 아래 결과를 보고 어떤 모델이 가장 적합한지 고르고 그 이유를 기술하시오
→ Random Forest



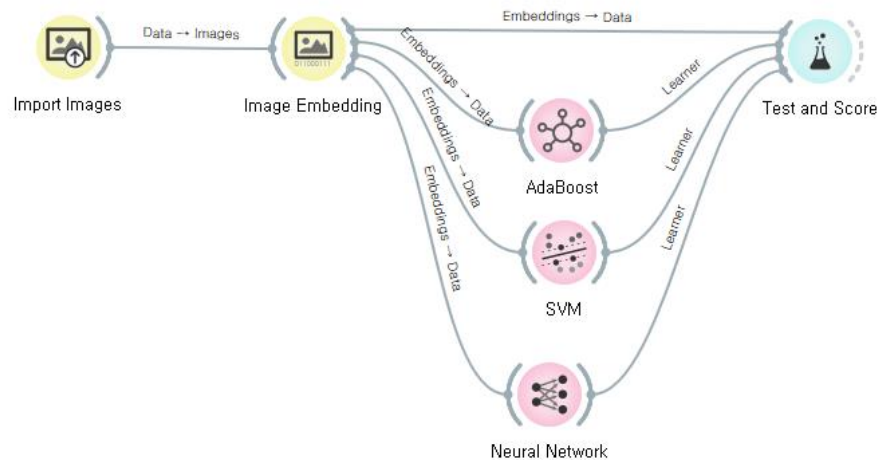
Test and Score - Orange

☐ Cross validation
 Number of folds: 10
☒ Stratified
☐ Cross validation by feature

☒ Random sampling
 Repeat train/test: 10
 Training set size: 80 %
☐ Stratified

Model	MSE	RMSE	MAE	R2
SVM	33.686	5.804	3.460	0.592
Random Forest	10.577	3.252	2.212	0.872
Linear Regression	23.599	4.858	3.311	0.714

16. 재활용품 이미지를 분류하는 머신을 만들고자 한다. 아래 결과를 보고 어떤 모델이 가장 적합한지 고르고 그 이유를 기술하시오.
→ Neural Network



Test and Score - Orange

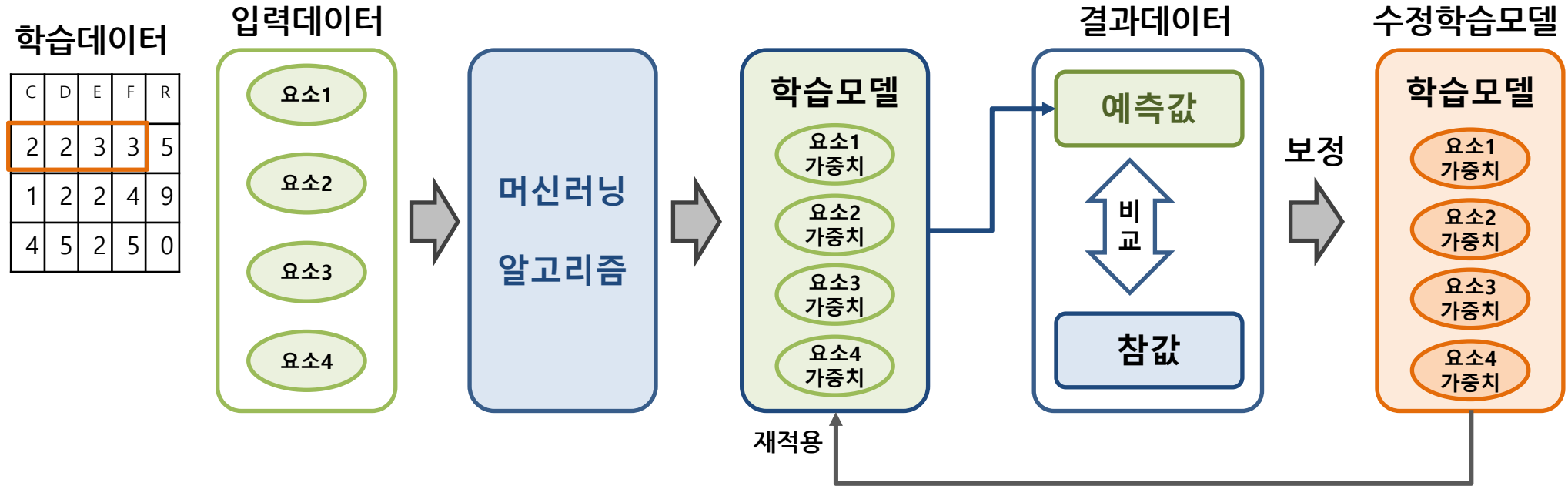
☒ Cross validation
 Number of folds: 5
☒ Stratified
☐ Cross validation by feature

☐ Random sampling
 Repeat train/test: 10
 Training set size: 80 %
☒ Stratified

Evaluation results for target (None, show average over

Model	AUC	CA	F1	Precision	Recall
SVM	0.988	0.835	0.810	0.871	0.835
Neural Network	0.981	0.876	0.872	0.875	0.876
AdaBoost	0.757	0.610	0.609	0.609	0.610

17. 기계가 지도학습으로 학습하는 원리/과정을 그림으로 나타내시오. (1개의 모델을 만드는 과정)



18. 머신러닝 학습방법 중 문제(Feature)와 정답(Label)이 있는 데이터(Training Set)를 컴퓨터에 학습시킨 후 새로 입력된 데이터(Test Set)를 분류하거나 예측하는 학습 방법을 무엇이라고 하는 가?

19. 다음 비지도 학습에 대한 설명 중 틀린 것은

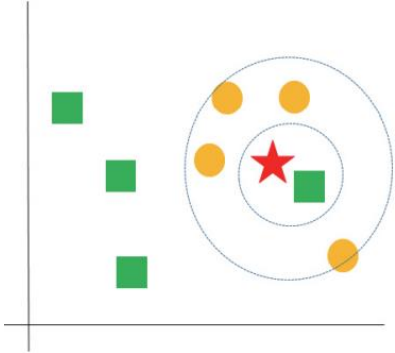
- 정답 레이블이 있음
- 입력된 데이터를 받아 스스로 데이터의 특징을 찾아 내는 것
- 데이터의 숨겨진 특징을 찾아내는 것에 목적
- 유사한 속성을 가진 값 끼리 그룹화하고 특징을 찾아낸다

20. 머신러닝 학습방법 중 '강화학습'에 대해서 설명하시오

21. 지도학습에 사용하는 알고리즘 중 3가지를 기술하시오

22. k-최근접이웃 알고리즘은 새로운 데이터를 어떻게 분류하는지 설명하시오

23. 아래 그림에서 별표를 새로운 데이터로 가정할 때 동그라미로 분류하게 하려면 K값을 어떻게 지정하면 되는 가?
(k-최근접이웃 알고리즘 적용을 가정함)



24. k-평균 알고리즘에 대한 설명 중 틀린 것은

- 유사한 특성을 가진 데이터 묶음을 클러스터라고 한다.
- K는 군집을 판단하기 위해 참조하는 데이터의 수를 의미한다.
- 정답 레이블을 사용하지 않는다.
- 비지도 학습 알고리즘 중 하나다.

25. 머신러닝 학습방법 중 '강화학습'에 대한 설명 중 틀린 것은

- 행동한 결과(처리 결과)에 대해 보상(평가) 알고리즘을 제공하여 스스로의 행동을 개선하는 방법이다
- 다른 답변을 탐색하지 않고도 훈련 데이터에서 직접 답을 적용한다
- 에이전트가 주어진 환경에서 어떤 행동을 취하고 이에 대한 보상을 얻으면서 학습 진행
- 지정한 작업을 수행하기 위해 훈련하는 프로그램을 에이전트라고 한다

26. 인공신경망에서 퍼셉트론(Perceptron)에 대해 기술하시오

27. 인공신경망에서 사용하는 활성화함수에는 어떤 것이 있는 가? (3개 이상 기술하시오)

28. 딥러닝에 대해서 기술 하시오.

29. 뉴런들의 연결 패턴을 구성한 전방향 인공 신경망으로 하위 레이어들에서 전처리를 수행하도록 설계한 신경망을 무엇이라고 하는 가?

30. '딥러닝'에 대한 설명 중 틀린 것은

- 딥러닝의 장점은 데이터의 특징을 단계별로 추상화를 높여 가면서 학습할 수 있다는 것이다
- 사람처럼(추상화 단계가 높은 특징을 사용해서 판단) 고차원적 인지활동을 수행할 수 있다.
- 원리가 간단하여 직접 구현하기 쉽고 정형화 된 데이터에서 타 알고리즘 보다 강점이 있다.
- 이미지나 자연어, 음성 등의 비정형화 된 대량의 데이터로부터 인식(Recognition)을 수행하는 문제영역에 잘 동작한다

실습 문제

1. 데이터마이닝 도구인 Orange3에서 모델 도구가 아닌 것은 ?



Linear Regression



Tree



Naive Bayes



Neural Network



Scatter Plot

2. 데이터마이닝 도구인 Orange3에서 데이터를 지도에 표시하고 싶을 때 어떤 Add-on 이 필요한가? (Add-on 추가 방법도 설명)

→ Options > Add-ons에서 Geo 애드온 추가

3. Orange3에서 데이터를 불러오는 도구가 아닌 것?



CSV File Import



Data Table



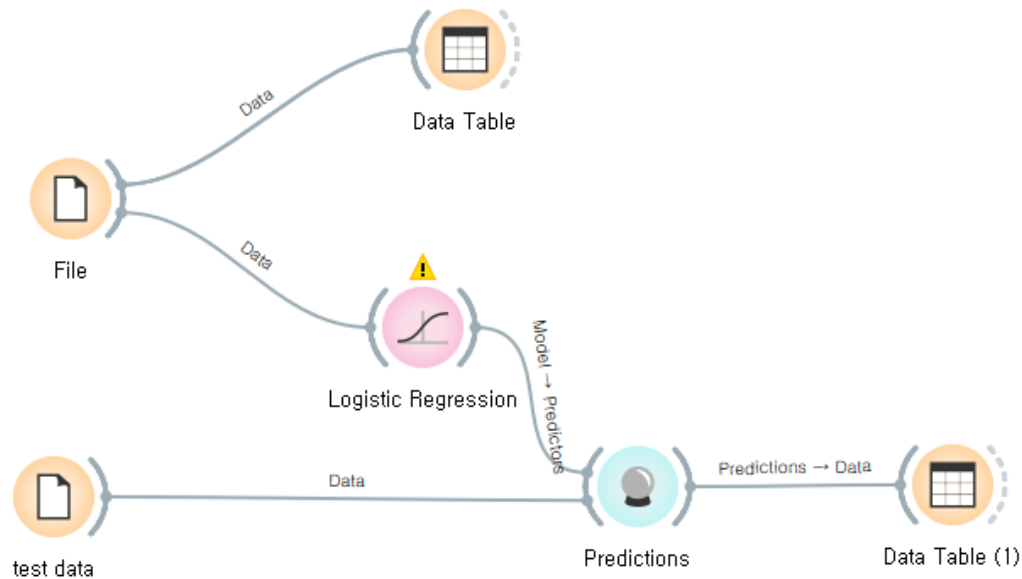
File



Datasets

4. 감귤평점을 자동 측정하는 모델을 Orange3로 만들려고 한다. 위젯 배치를 그림으로 도시하시오.

- 데이터 불러오기 : File '감귤평점훈련', '감귤평점테스트'
- 데이터 보기 : Data Table
- 데이터 학습 모델 : Logistic Regression
- 모델 평가 : Predictions



5. Orange3에서 3가지 학습 알고리즘을 이용하여 모델을 만들었다. Confusion Matrix에서 다음과 같은 결과를 얻었을 때 어떤 모델이 가장 적합하다고 할 수 있는 가? (그 이유도 기술) → Naïve Bayes

Confusion Matrix - Orange

Learners		Predicted		
		0	1	Σ
Actual	0	104	28	132
	1	34	125	159
	Σ	138	153	291

Learners		Predicted		
		0	1	Σ
Actual	0	107	25	132
	1	26	133	159
	Σ	133	158	291

Learners		Predicted		
		0	1	Σ
Actual	0	108	24	132
	1	24	135	159
	Σ	132	159	291

6. 다음은 카메라로 촬영된 사람의 감정을 인식하는 블록프로그램이다. 빈 곳에 들어가야 하는 블록은 ?



답은 실습예제 유인물 (깃허브에도 있음)에서 찾아서 확인하실 것

7. 카메라로 사람얼굴을 촬영하여 나이, 감정, 성별 등을 판단해 주는 블록을 사용하려고 한다.
어떤 스프라이트 확장이 필요한가?

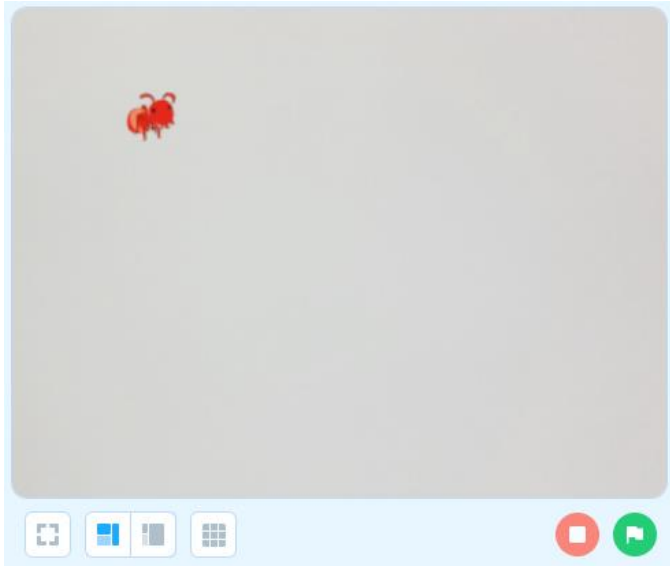
- 1) 인식서비스
- 2) Translation
- 3) Video Sensing
- 4) Text to Speech

8. 다음 블록프로그램은 한국어를 입력하면 영어로 대답해 주는 프로그램이다.
다음 빈 곳에 들어갈 블록은?

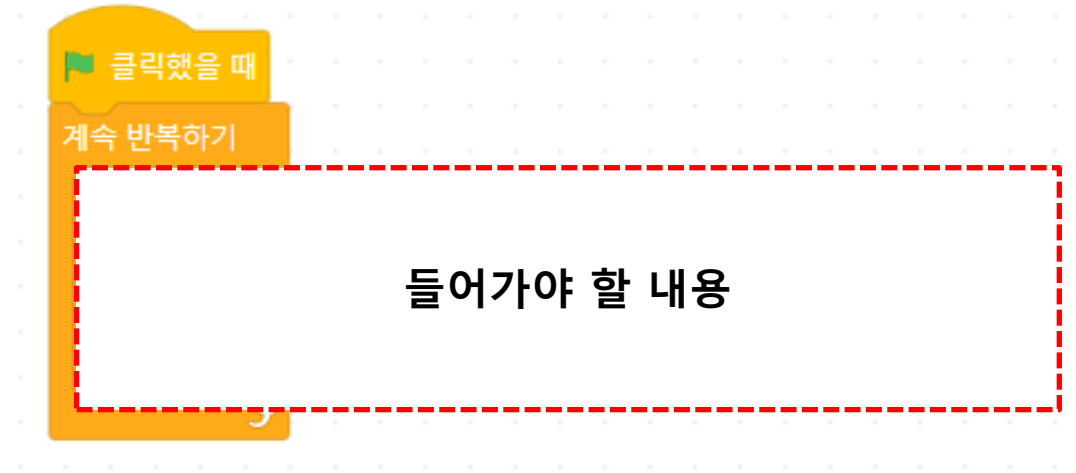
답은 실습예제 유인물 (깃허브에도 있음)에서 찾아서 확인하실 것



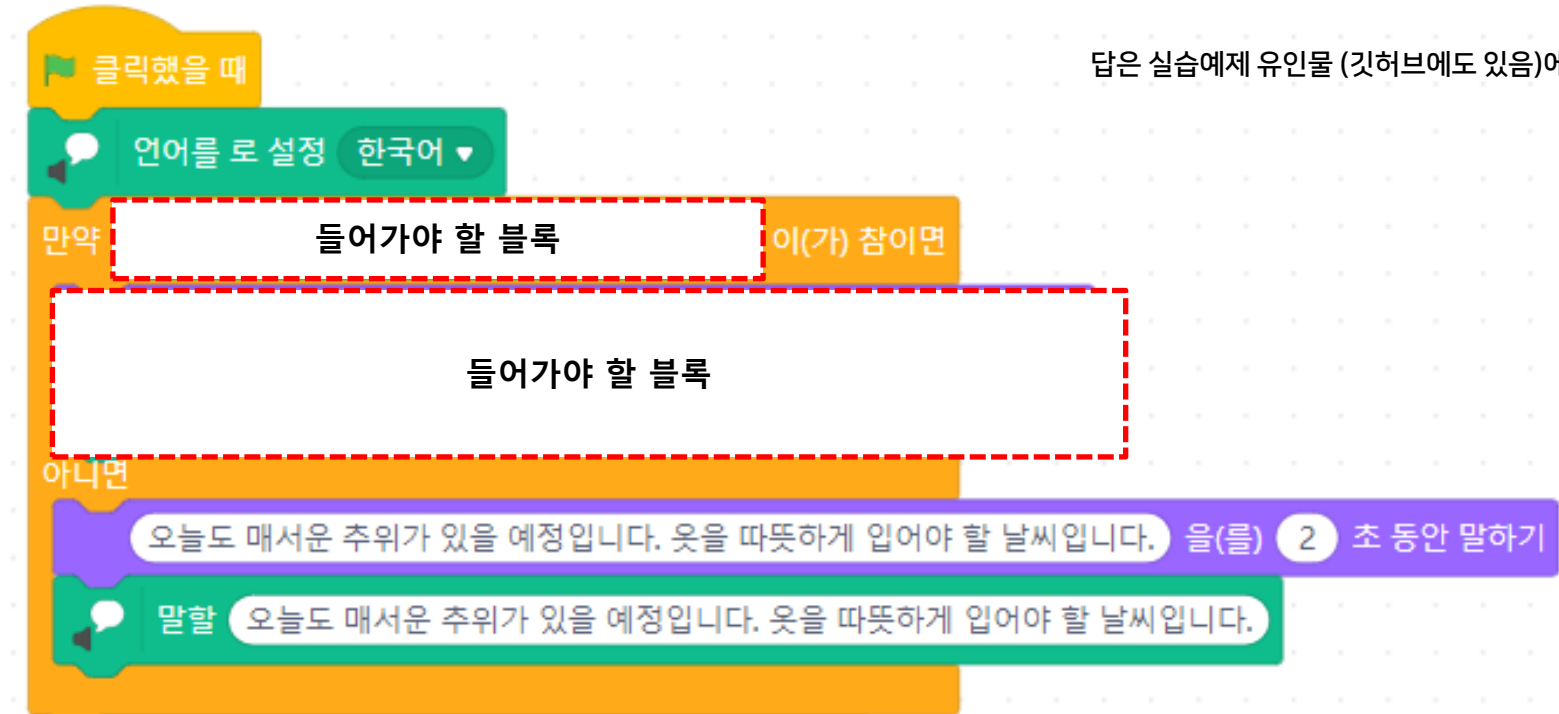
9. 다음 블록프로그램은 카메라가 촬영하는 화면을 인식해서 움직임이 있는 부분이 개미에 닿으면 개미가 전진하는 프로그램이다. 손을 움직여 개미에 닿게 했을 때 개미가 앞으로 이동하는 블록프로그래밍을 하시오.



답은 실습예제 유인물 (깃허브에도 있음)에서 찾아서 확인하실 것



10. 다음은 날씨를 안내하는 프로그램이다. 서울 최저 온도가 18도 이상이면 '오늘의 날씨는 따뜻하게 시작하겠습니다.'를 출력하고자 할 때 빈 곳에 들어갈 블록을 그리시오.
(서울 = Seoul,KR)



답은 실습예제 유인물 (깃허브에도 있음)에서 찾아서 확인하실 것