

교육용프로그래밍언어기초(스크래치)

8주차 - 2교시

머신러닝



 머신러닝과 머신러닝 모델 생성 과정

 머신러닝 학습방법



학습목표

- ① 머신러닝과 머신러닝 모델 생성과정을 이해할 수 있다.
- ② 머신러닝의 지도학습을 이해할 수 있다.
- ③ 머신러닝의 비지도학습을 이해할 수 있다.
- ④ 머신러닝의 강화학습을 이해할 수 있다.

1

머신러닝과 머신러닝 모델 생성 과정



1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

1 아서 사무엘의 정의

- 머신러닝용어는 1959년에 아서 사무엘이 학술지 <IBM Journal of Research and Development>에 기고한 논문에서 처음 사용함

머신러닝(Machine Learning)

컴퓨터가 명시적으로 프로그램되지 않고도 학습할 수 있도록 하는 연구 분야

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

1 아서 사무엘의 정의

“머신러닝은 컴퓨터가 명시적으로 프로그램되지 않고도 학습할 수 있도록 하는 연구 분야를 말합니다.”

(Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.)

-아서 사무엘-

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

1 아서 사무엘의 정의

- 일반적인 프로그래밍 작업
 - 입력값에 따라 원하는 결과값이 출력되도록 **사람이 내부 동작을 작성함**
- 기계학습
 - 사람이 컴퓨터에게 입력값과 결과값만 충분히 전달해 주면 **컴퓨터가 스스로** 입력값과 결과값의 관계를 만족시키는 **내부 동작을 찾아냄**

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

1 아서 사무엘의 정의

학습 데이터
(훈련 데이터)

내부 동작을 만들 때 사용한 데이터
(입력값과 결과값)

시험 데이터

만들어진 내부 동작의 성능을 평가할 때
사용하는 데이터

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

1 아서 사무엘의 정의

- 머신러닝은 여러 개의 입력값과 결과값을 컴퓨터에 제공하기만 하면 이 데이터를 바탕으로 컴퓨터가 스스로 내부 동작을 만들어 냄

Point

사람이 할 일은 단지 학습 데이터를 공급하는 일 뿐

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

2 카네기 멜론 대학교 교수인 토미 미첼의 정의

**“만약 컴퓨터 프로그래밍이 어떤 작업 T를 수행할 때, 경험 E를
통해 성능이 향상된다면(성능 측정 방법은 P), 그 컴퓨터
프로그램은 작업 T를 수행할 때
방법 P로 측정되는 성능 향상을 경험 E로부터 학습한다고 말할
수 있습니다.”**

(A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E)

-토미 미첼-

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

2 카네기 멜론 대학교 교수인 톰 미첼의 정의

- 만약 머신러닝 방식으로 고양이 판별 프로그램을 개발한다면, 고양이 판별 정확도를 높이기 위해 프로그램에 **수많은 고양이 사진들을 입력**해 훈련시켜야 함

1 머신러닝과 머신러닝 모델 생성 과정

1 머신러닝의 개념

2 카네기 멜론 대학교 교수인 톰 미첼의 정의

- 〈고양이 판별 프로그램 예를 활용한 톰 미첼의 머신러닝 정의 해석〉

작업 T, 성능 측정 P, 경험 E	해석
T : 고양이 판별하기	〈고양이 판별하기〉 작업을 수행할 때 ‘고양이 사진 입력’으로 ‘고양이 판별 정확도’가 향상된다면, 이 프로그램은 〈고양이 판별하기〉 작업을 수행할 때 ‘고양이 판별 정확도’의 성능을 향상시키기 위해 ‘고양이 사진 입력’이란 경험으로부터 학습함
P : 고양이 판별 정확도	
E : 고양이 사진입력	

1 머신러닝과 머신러닝 모델 생성 과정

2 머신러닝 모델의 생성 과정

- 머신러닝으로 문제를 해결하기 위해서는 문제 해결에 적합한 머신러닝 모델을 생성해야 함

문제 속의 데이터를 잘 설명할 수 있는
머신러닝 모델을 가정



모델로부터 학습 데이터에 최적화된
구체적인 함수를 찾음

학습
(Learning)

- 이렇게 학습된 함수를 실제 문제에 적용하면 됨

1 머신러닝과 머신러닝 모델 생성 과정

2 머신러닝 모델의 생성 과정

- 시험 공부 시간으로 시험 성적을 예상하는 문제를 머신러닝으로 해결하고자 할 때의 머신러닝모델 생성 과정

머신러닝 모델 가정

시험 공부 시간으로 시험 성적을 예상할 수 있을까?
시험공부를 많이 하면 성적도 오를 것이라 생각해.
직선과 같은 모델을 가질 것이라 가정하자

머신러닝 모델 학습

어떤 직선이 시험 공부 시간과 시험 성적 간의 관계를 가장 잘 나타낼까?
친구들의 시험 공부 시간과 시험점수를 모아보자.
그리고 그 데이터를 학습데이터로 사용하여 가장 적절한 직선을 찾아보자

머신러닝 모델 적용

이제 시험 공부 시간을 입력하여 점수를 예측해보자

1 머신러닝과 머신러닝 모델 생성 과정

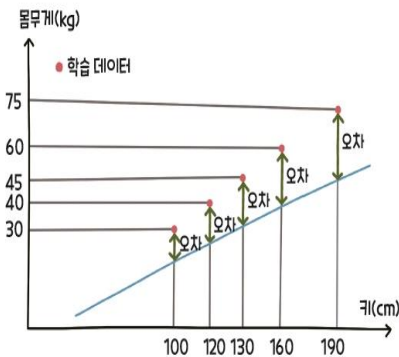
3 머신러닝 구현 과정 예제

- 사람의 키를 입력했을 때 몸무게를 추측하는 작업을 머신러닝으로 구현하면?
 - ① 일반적으로 키가 커지면 몸무게도 늘어날 것이라 가정해, [몸무게 = $(a \times \text{키}) + b$]와 같은 직선의 방정식을 만들어 머신러닝모델로 가정
 - ② 머신러닝모델로 가정한 직선의 방정식은 아직 기울기 a 와 y 절편 b 의 값이 결정되지 않은 상태로, 학습 데이터를 확보하여 **최적화된 직선**을 구함

학습 데이터와 최적화된 직선을 구한다는 것은 학습 데이터와 오차가 가장 적은 직선의 기울기와 y 절편을 구한다는 의미

3 머신러닝 구현 과정 예제

- 사람의 키를 입력했을 때 몸무게를 추측하는 작업을 머신러닝으로 구현하면?

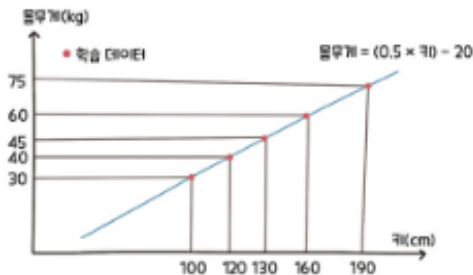


〈학습 데이터와 직선간의 오차〉

학습 결과, 직선의 방정식 [몸무게 = $(0.5 \times \text{키}) - 20$] 이 해당 학습 데이터에 최적화된 함수임

③ 머신러닝 구현 과정 예제

- 사람의 키를 입력했을 때 몸무게를 추측하는 작업을 머신러닝으로 구현하면?



〈몸무게와 키의 상관관계에 최적화된 직선의 방정식〉

③ 최적화된 머신러닝모델을 실제 문제에 적용해 확인

- 학습 데이터에는 없던 키 180cm를 방정식에 대입하면 예측되는 몸무게는 $[70\text{kg} = (0.5 \times 180) - 20]$ 이 나옴



2

머신러닝 학습방법





2

머신러닝 학습방법

1 지도학습 (Supervised Learning)

지도학습

입력값에 대한 정답이나 결과값을 알고 있는 학습 데이터를 활용하여 머신러닝모델을 학습시키는 방식

입력값에 대한 정답 또는 결과값을
레이블 (Label)이라고 함

1 지도학습 (Supervised Learning)

- 〈구분된 학습 데이터를 활용하는 지도학습의 예〉



1 지도학습(Supervised Learning)

- 주로 분류 또는 회귀와 같은 작업을 수행할 머신러닝모델 학습에 사용

분류작업 (Classification)

어떤 입력 데이터가 들어오더라도 학습에 사용한 레이블 중 하나로 결과값을 결정하는 작업

회귀작업 (Regression)

입력 데이터에 대한 결과값으로 학습에 사용한 레이블 외의 값이 나올 수 있는 작업

1 지도학습(Supervised Learning)

- 〈지도학습을 적용한 머신러닝 모델의 작업〉

작업	내용
분류	입력값에 대한 결과값이 정해진 레이블 중 하나로 결정되는 작업을 의미
회귀	입력값에 대한 결과값이 학습에 사용된 레이블 외의 값도 나올 수 있는 작업을 의미

1 지도학습(Supervised Learning)

- 〈지도학습을 적용한 머신러닝 모델의 작업〉

필기체(손글씨) 인식은 대표적인 **분류 작업**의 예

- 0부터 9까지의 숫자 이미지 레이블을 학습 데이터로 사용
- 어떤 이미지라도(심지어는 숫자 이미지가 아니더라도)
0부터 9까지의 레이블 중 하나로 결과값을 결정함

2

머신러닝 학습방법

1 지도학습(Supervised Learning)

- 〈지도학습을 적용한 머신러닝 모델의 작업〉

몸무게 예측 작업은 회귀 작업의 예

- 여러 명의 키와 그에 대응하는 몸무게를 학습 데이터로 사용하는데, 입력된 키에 대한 몸무게를 결과값으로 출력하기 때문에 몸무게가 레이블이 됨
- 다양한 키에 대해서 레이블에 포함되지 않은 몸무게도 결과값으로 출력 가능



2

머신러닝 학습방법

2 비지도학습 (Unsupervised Learning)

비지도학습

입력값에 대한 정답이나 결과값이 없는, 즉 레이블이 없는 데이터를 사용하여 머신러닝모델을 학습시키는 방식

대표적인 작업으로 **군집 작업**이 있음

② 비지도학습(Unsupervised Learning)

- 군집(Clustering) 작업
 - 데이터를 구분하여 비슷한 집단으로 묶는 작업
 - 데이터를 구분한다는 점에서 군집 작업과 분류 작업은 비슷하지만, 정해진 레이블 없이 데이터 간에 존재하는 비슷한 속성을 기준으로 데이터를 묶는다는 것이 차이점
 - 군집 작업에 필요한 기준 속성은 데이터를 기반으로 머신러닝모델이 스스로 찾아냄

② 비지도학습(Unsupervised Learning)

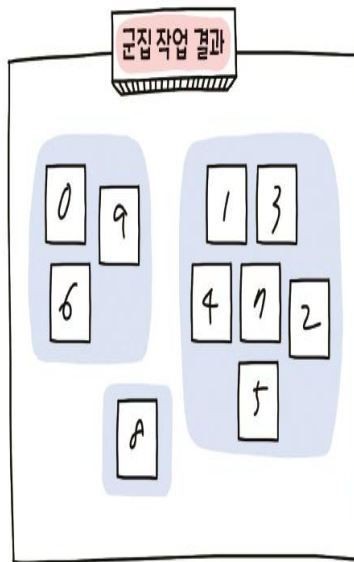
- 군집(Clustering) 작업

숫자 필기체 이미지를 대상으로 군집 작업을 수행한다면?

- 0~9까지 구분된 레이블과는 상관없이 숫자 필기체 이미지 자체를 학습 데이터로 사용하게 되는데, 머신러닝모델은 스스로 숫자 필기체 이미지에 존재하는 속성을 찾아내고, 유사한 속성을 갖는 이미지들끼리 묶는 작업을 수행함

② 비지도학습(Unsupervised Learning)

- 군집(Clustering) 작업
 - 군집 작업에 사용된 데이터 속성은 각 이미지에 포함된 원 모양의 개수



〈머신러닝 모델의 군집 작업 (예: 숫자 이미지의 군집)〉

② 비지도학습(Unsupervised Learning)

- 군집(Clustering) 작업
 - 군집 작업은 **마케팅 분야**에서 많이 응용되고 있음

예

채팅 사이트를 운영하는 기업이 사람들의 대화 내용을 기반으로 군집 작업을 수행해 취미를 속성으로 묶었다면, 집단별로 인구통계학적 특성을 파악해 적절한 취미용품을 광고할 수 있음

검색 포털 사이트 내 유사한 뉴스 그룹핑, 소셜 네트워크 서비스(SNS)의 사용자 간 친밀도 분석, 보안 분야의 이상 행위 탐지, 카드사의 이상 거래 탐지 등에 다양하게 활용되고 있음

2

머신러닝 학습방법

③ 강화학습(Reinforcement Learning)

강화학습

입력값에 대한 정답이나 결과값 대신 어떤 일을 잘했을 때 보상을 주는 것으로 머신러닝모델을 학습시키는 방식

2

머신러닝 학습방법

③ 강화학습(Reinforcement Learning)

- 강화학습의 핵심 목표
 - 시행착오를 통해 **보상의 총합이 최대가 되는 일련의 행동**을 찾는 것
 - 나중에 더 큰 이익을 기대하며 지금의 손해를 감수하는 식의 전략적 행동을 탐색할 수 있다는 의미와 같음

③ 강화학습(Reinforcement Learning)

- 강화학습의 핵심 목표

사업 분석 분야

게임 전략,
금융시장의 투자 전략,
광고 노출 전략 등

로보틱스분야

로봇 팔 제어,
이족보행 제어 등

자율주행 분야

③ 강화학습 (Reinforcement Learning)

- 〈강화학습을 통한 게임 제어 활용 예〉



(a) 슈퍼마리오(<https://www.youtube.com/watch?v=WzxmH1Cx2Yg>)



(b) 쿠키런(<https://www.youtube.com/watch?v=exXD6wJLJ6s>)

교육용프로그래밍언어기초(스크래치)

Next

인공신경망과 딥러닝

