Chapter 02. 금융 AI 모델의 동작원리

■ 핀테크 AI 모델은 어떤 원리로 데이터를 빠르고 정확하게 분석하나?



2-1. AI는 "데이터"를 가지고 판단을 스스로 합니다

■ 금융 데이터, 금융 정보, 금융 지식을 구분할 수 있나요?

데이터 (Data)

- 자료(資料)라고도 하며, 수, 영상, 단어 등의 형태로 된 의미 단위임
- 보통 연구나 조사 등의 바탕이 되는 재료를 말함

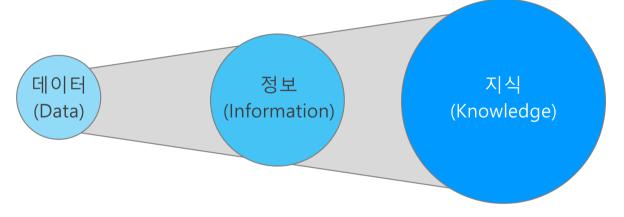
내일 날씨 예보와 같이 데이터를 처리하여 나온 결과를 '정보(Information)'이라고 함

즉, 정보란 사용자 의사결정에 도움이 될 만한 잘 정리된 데이터라고 할 수 있음

지식(Knowledge)은 이러한 정보를 바탕으로 의사결정을 돕기 위한 규칙으로 변환한 것임

2-2. AI 모델의 원리

■ 데이터, 정보, 지식의 구분



"S전자 2017년도 매출액이 100조원입니다." "S전자 2021년 실적은 양호한 편입니다." "주가가 오를 가능성이 높으니 <mark>지금</mark> 사는 것이 유리하다."









2-2. AI 모델의 원리

■ 기계학습(Machine Learning)으로 데이터에서 정보를 발견하는 방법

"사람의 개입 없이 머신 스스로 학습"

즉, 학습 데이터(Training Data)로 스스로 학습해서 정보를 발견한다.

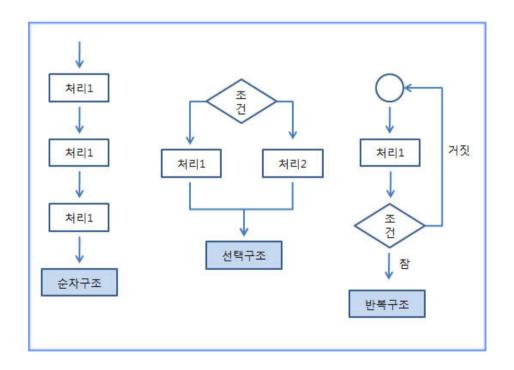
(예시: 스팸메일의 분류, 주가예측, 마케팅의 고객 세그멘테이션)

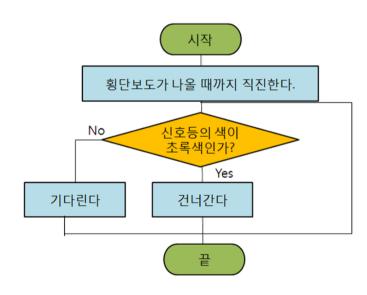


<머신러닝의 개념>

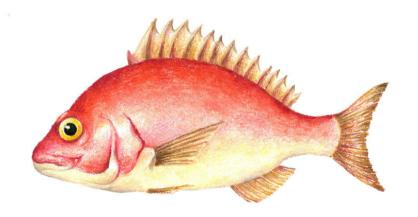
2-3. AI 모델과 코딩(프로그래밍)

■ AI 모델: IF(조건문) 없이 판단이 가능한가?





■ AI에게 물고기를 구분하는 일을 시켰을 경우



4	Α	В	С	D	E	F
4			C	D		
1	Length	Weight				
2	23.2	242				
3	24	290				
4	23.9	340				
5	26.3	363				
6	26.5	430				
7	26.8	450				
8	26.8	500				
9	27.6	390				
10	27.6	450				
11	28.5	500				
12	28.4	475				
13	28.7	500				
14	29.1	500				
15	29.5	340				
16	29.4	600				



	Α	В	С	D	Е	F
1	Length	Weight				
2	23.2	242		1		
3	24	290		1		
4	23.9	340		1		
5	26.3	363		1		
6	26.5	430		1		
7	26.8	450		1		
8	26.8	500		1		
9	27.6	390		1		
10	27.6	450		1		
11	28.5	500		1		
12	28.4	475		1		
13	28.7	500		1		
14	29.1	500		1		
15	29.5	340		1		
16	29.4	600				



	Α	В	С	D	Е	F
1	Length	Weight	target			4
2	23.2	242	1			1
3	24	290	1			1
4	23.9	340	1			1
5	26.3	363	1			1
6	26.5	430	1			1
7	26.8	450	1			1 1
8	26.8	500	1			
9	27.6	390	1			1
10	27.6	450	1			1 1
11	28.5	500	1			
12	28.4	475	1			1
13	28.7	500	1			1
14	29.1	500	1			1
15	29.5	340	1			1
16	29.4	600	1			



학습정리 #2차

■ 인공지능 모델의 개발 프로세스

- 인공지능 모델에서는 판단에 해당하는 "IF"(조건)를 사용자가 직접 입력하지 않음
 - : 데이터 수집 → 인공지능 모델(회귀, 분류 등) 구성 → 값의 예측 또는 Class의 자동 분류 → 검증

■ 인공지능 모델의 학습 원리

- 데이터에서 어떤 특성(Feature)를 발견하는 과정
- : 전문가가 아니더라도 잘 정리된 데이터를 통해서 우리가 원하는 답을 찾을 수 있음
- 우수고객/보통고객, 정상거래/이상거래 등 특이한 패턴을 학습을 통해서 발견할 수 있음

■ 기계학습(Machine Learning)이란

- 기계학습은 학습 데이터(Training Data)로 스스로 학습해서 정보를 발견하는 과정임
- 기계학습에서는 모델을 도출한 뒤에 이를 평가(Test)하는 과정이 필요하며, 이러한 평가는 "정확도"로 측정됨