

강원지역혁신플랫폼

# 기계학습

Machine Learning



100%

추론 및 기계학습 활용의 개념





## ▶ 학습목표

📁 추론 및 기계학습 활용의 개념을  
설명할 수 있습니다.





# 01 | 추론 및 기계학습



## 추론 및 기계학습

### △ 추론 및 기계학습

- ◆ 딥러닝으로 인한 인공지능의 발전은 인지, 학습, 추론, 행동과 같은 인간 지능의 전 과정에 걸쳐 혁신적인 진화를 만들어내고 있음
  - 시각, 청각과 같은 감각기관에 해당하는 인지 지능에서부터 스스로 지능을 발전시키는 학습
  - 새로운 상황을 추론하고 행동하는 단계에 이르기까지 다양한 분야의 연구가 동시다발적으로 빠르게 발전되고 있음
  - 추론 및 기계학습은 인간의 사고능력을 모방하는 기술들로 추론, 인공신경망, 강화 학습, 딥러닝, 인지공학 등이 있음



## 02 | 기계학습



### 기계학습(Machine Learning)

△ 기계학습은 인공지능의 한 분야로 머신러닝이라고도 부름

◆ 데이터 분석을 위한 모델 생성을 자동화하여 소프트웨어가 데이터를 바탕으로 학습하고 패턴을 찾아냄

‣ 사람의 개입을 최소화하고 빠르게 의사결정을 내릴 수 있도록 지원함

– 기계가 명시적으로 프로그래밍 되지 않은 상태로 알고리즘을 사용하여  
데이터로부터 학습해서 실행하는 것임

– 즉, 명시적으로 프로그래밍 되지 않고도 학습할 수 있는 능력을 가진 알고리즘

– 데이터를 통해서 학습하기 위한 특정 비즈니스 규칙을 제공할 필요가 없음



## 02 | 기계학습



### 기계학습(Machine Learning)

△ 기계학습은 컴퓨터 과학 중 인공지능의 한 분야로 패턴인식과 컴퓨터 학습 이론의 연구로부터 진화한 분야

◆ 경험적 데이터를 기반으로 학습을 하고 예측을 수행하고 스스로의 성능을 향상시키는 시스템

➢ 이를 위한 알고리즘을 연구하고 구축하는 기술

➢ 입력 데이터를 기반으로 예측이나 결정을 이끌어내기 위해  
특정한 모델을 구축하는 방식을 취함



## 02 | 기계학습



### 기계학습(Machine Learning)

- ⌘ 기계학습 기술은 특정한 과제를 수행하도록 프로그래밍하지 않아도 컴퓨터가 학습할 수 있다는 이론과 데이터 패턴 인식이 어우러져 탄생
  - ◆ 컴퓨터가 데이터를 통해 스스로 학습할 수 있음
    - 새로운 데이터에 노출됨에 따라 독립적으로 최적화를 수행한다는 점에서 기계학습에서는 반복적 측면이 중요함
      - ─ 이전 연산 결과를 학습하여 믿을 수 있는 의사 결정 및 결과를 반복적으로 산출하기 때문



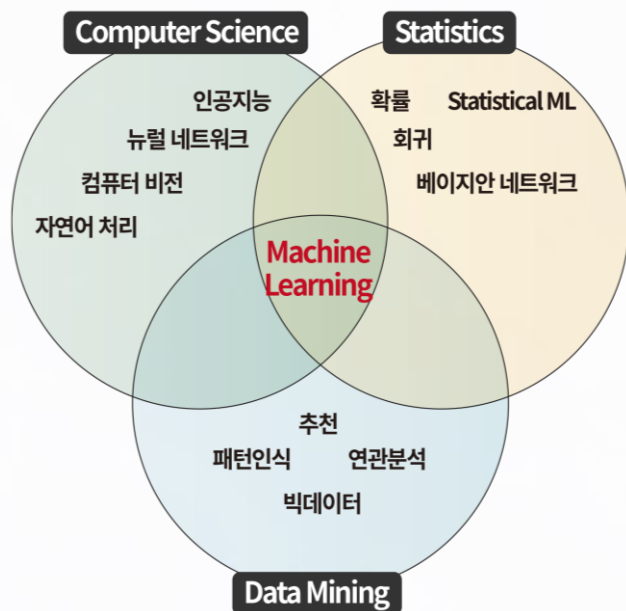
## 02 | 기계학습



### 기계학습(Machine Learning)

△ 기계학습에 사용되는 **리소스**(Resource) **분류**는 아래 그림과 같다.

- ◆ 컴퓨터 사이언스(Computer Science)
- ◆ 통계(Statistics)
- ◆ 데이터 마이닝(Data Mining)





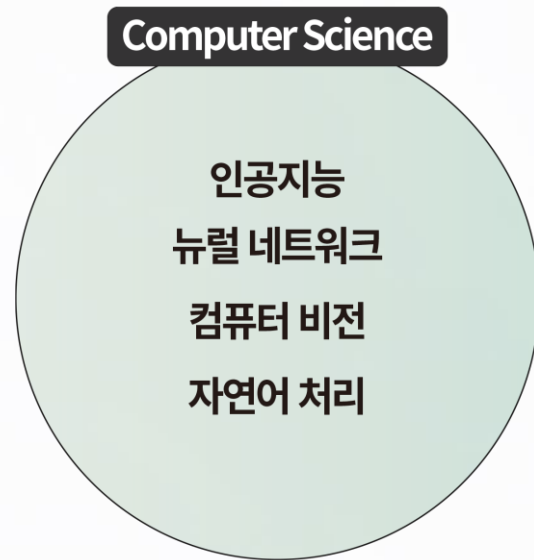
## 02 | 기계학습



### 기계학습(Machine Learning)

#### ◆ 컴퓨터 사이언스(Computer Science)

- ▶ 컴퓨터 사이언스는 **과학 기술의 한 분야**
- ▶ **자동적 수단**에 의하여 **수행되는 데이터 처리에** **관련하는 방법과 기술**의 한 분야
- ▶ 인공지능, 뉴럴 네트워크, 컴퓨터 비전, 자연어 처리 등이 있음







## 02 | 기계학습



### 기계학습(Machine Learning)

#### ◆ 통계(Statistics)

- ▶ 관측 대상이 되는 표본(Sample) 집합으로부터 얻어지는 여러 가지 측정값
- ▶ 대량의 관측 값에 대해 여러 가지 연산을 행하여  
확률, 회귀, 베이지안 네트워크 등 여러 가지 형태로 통계 값을 찾아낼 수 있음





## 02 | 기계학습



### 기계학습(Machine Learning)

#### ◆ 데이터 마이닝(Data Mining)

- ▶ 데이터 베이스 내에서 어떠한 방법(추론, 패턴 인식, 연관분석, 유사성 등)에 의해 관심 있는 지식을 발견하는 과정
- ▶ 대용량의 데이터 속에서 유용한 정보를 발견하는 과정
- ▶ 기대했던 정보뿐만 아니라 기대하지 못했던 정보를 찾을 수 있는 기술
- ▶ 정보의 연관성을 파악함으로써 가치 있는 정보를 만들어 의사 결정에 적용함으로써 이익을 극대화하거나 위험을 회피하는 방법에 사용될 수 있음





## 02 | 기계학습



### 기계학습(Machine Learning)

△ 기계학습에 대한 관심은 데이터 마이닝 등과 같은 기술의 발전에서 찾아볼 수 있음

◆ 기계학습에 대한 지속적인 관심을 불러 일으키는 요소는 다음과 같음

- ▶ 사용 가능한 데이터의 볼륨과 다양성의 증가
- ▶ 분석 비용의 감소
- ▶ 강력해진 분석 기술
- ▶ 저렴한 스토리지 비용 등



## 02 | 기계학습



### 기계학습(Machine Learning)

- ⌘ 기계학습은 아무리 규모가 큰 데이터라도 분석 모델을 자동으로 빠르게 생성함으로써 복잡한 분석에서 정확한 결과를 도출할 수 있음
  - ◆ 기업들은 분석 결과를 이용하여 수익성이 높은 기회를 찾아내거나 미지의 위험을 회피하는 등 인사이트를 획득할 수 있음
  - 기계학습은 알고리즘을 이용해 연계성을 찾아내는 모델을 구축함으로써 조직은 사람의 개입 없이도 더 나은 의사 결정을 내릴 수 있음



## 03 | 추론



### 추론 (Inference)

△ 인간의 두뇌는 ‘학습’과 ‘추론’으로 구분됨

◆ 다양한 지식과 정보를 머리 속에 축적하는 ‘학습(또는 훈련)’

➢ 그 지식을 기반으로 새로운 정보에 대한 답을 스스로 도출해내는 ‘추론’







## 03 | 추론



### 추론 (Inference)

- ⚡ 딥러닝 또한 인간의 학습 과정을 그대로 옮겨와 그 나름의 ‘훈련’ 과 ‘추론’을 거쳐 인공지능을 구현하게 됨
  - ◆ 훈련 과정이 없으면 추론도 할 수 없음
    - 우리가 배운 지식을 토대로 사고하는 것처럼 훈련을 거치는 과정은 필수임
      - ▬ 우리가 고학년이 될수록 혼자 책을 읽을 수 있는 것처럼,  
심층 딥러닝도 지도가 없이도 정답을 도출해 낼 수 있게 됨



## 03 | 추론



### 추론 (Inference)

△ 오늘날 AI는 더 많은 데이터와 보다 빠른 처리능력 그리고 더 강력한 알고리즘이 결합되어 더욱 빠른 속도로 확산되고 있음

◆ 지금까지 AI는 모델 개발 및 훈련에 정통한 AI 전문가(분석 전문가)들의 전유물

▶ 하지만, 지금은 AI 중심이 이니셔티브(문제를 해결하기 위한 계획)가 비즈니스의 전 범위로 확장되면서 상황은 달라지고 있음



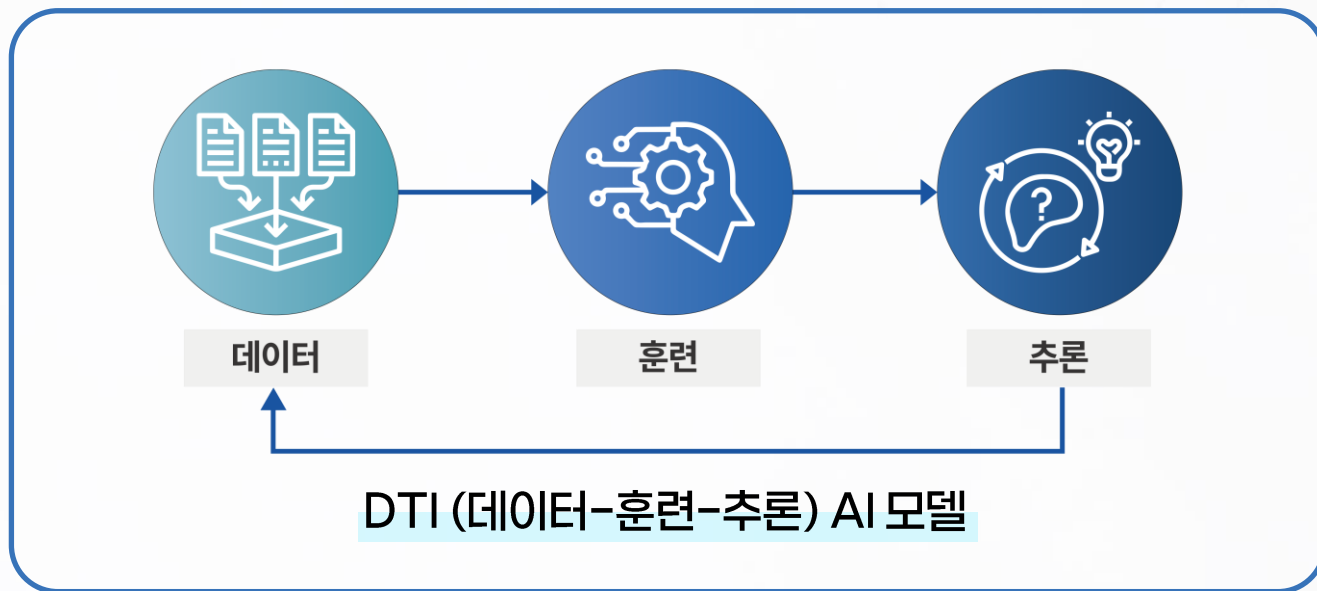
## 03 | 추론

### 추론 (Inference)

> 비즈니스의 전 범위로 확장에는 가치 기반의 AI 활동 프레임워크(DTI)에 있음

— 데이터(Data) → 훈련(Train) → 추론(Inference) AI 모델

→ 끊임없이 상호 작용하는 세 개의 단계가 연속 루프를 이루는 형태임





## 03 | 추론

### 추론 (Inference)

#### 1 데이터 단계

- ◆ AI 전문가들은 이 세 단계 중에서 ‘데이터 단계’에 가장 많은 시간이 소요된다고 지적함
  - AI 모델 훈련에 사용할 데이터를 준비하는 작업에 많은 시간이 필요
  - 오랜 격언대로 ‘쓰레기가 들어가면 쓰레기가 나오는 법’
  - 데이터 품질은 AI를 성공적으로 구축하는데 가장 큰 과제임
- 일반적으로 정확한 AI 모델을 훈련하는 데는 많은 데이터가 필요하다고 알려져 있음
  - ─ 정상적인 시스템 작동을 위해 많은 데이터가 필요하지만,  
실제로 필요한 것은 이상 또는 심각한 장애 조건의 데이터일 것임



## 03 | 추론



### 추론 (Inference)



#### 1 데이터 단계

- ◆ 데이터는 각기 다른 여러 소스로부터 얻고, 가공되지 않은 형식임
  - 즉, 텍스트, 이미지, 사운드 또는 원시 상태의 수치일 수 있음
  - 데이터 사이언스 팀은 데이터를 수집 및 정리하여 해당 프레임워크 및 분석 환경에서 사용하기에 적합한 형식으로 만드는 데 많은 시간을 보냄
  - 기업의 데이터는 대개 파편화되어 있거나 여러 위치에 분산되어 있음





## 03 | 추론



### 추론 (Inference)

#### 1 데이터 단계

◆ 데이터는 적시성이 매우 중요함

➢ 환경적 조건이 급변할 수 있으므로 AI 모델에 최신 데이터를 공급해야  
계속해서 비즈니스 가치가 창출될 수 있음

➢ 새로운 데이터가 없으면 모델의 가치는 떨어지게 됨

➢ 데이터를 적절하게 유지하려면 모델 훈련을 위한  
기본 데이터 세트를 끊임없이 업데이트하는 계획을 마련해야 함

— 즉, 가치 있는 인사이트를 발굴하기 위해서는  
잘 정리된 유의미하고 신선한 데이터가 지속적으로 필요함



## 03 | 추론



### 추론 (Inference)

#### 2 훈련 단계

- ◆ 훈련 단계는 이전 단계의 데이터를 사용하여 모델을 만드는 반복적인 프로세스임
  - 훈련 단계는 데이터가 AI 모델로 변화하는 인공지능의 마법이 일어남
    - ─ 이 모델에서는 현실 세계의 비슷한 데이터에 기초하여 미래를 예측할 수 있음
    - ─ 딥러닝의 심층 신경망 모델이 이러한 방식으로 문제를 해결하는 것이 가능해졌는데, 그것은 GPU(Graphics Processing Unit)가 등장한 덕분임
  - 모델 훈련을 위해 성능을 가속하면 큰 보탬이 되지만, 속도와 정확성이 균형을 이뤄야 시장 출시 일정 측면에서 중요한 가치를 실현할 수 있음



## 03 | 추론

### 추론 (Inference)

#### 2 훈련 단계

- ◆ 훈련 단계에서 가장 많은 시간이 소요되는 것은 모델에 대한 하이퍼 파라미터를 설정 및 재설정하는 작업
  - 하이퍼 파라미터는 데이터 사이언티스트가 모델 훈련을 시작하기 전에 해당 모델을 위해 선택하는 값을 말함
    - ─ 최신 모델 중에는 하이퍼 파라미터가 수백 개에 달할 수 있음
  - AI 모델은 기존 코드와 달리 새로운 데이터로 재훈련하지 않으면 차츰 기본 데이터에서 멀어지게 됨
    - ─ 기존 모델을 계속 재훈련해야 유의미하고 유용한 모델로 유지할 수 있음



## 03 | 추론

### 추론 (Inference)

#### 3 추론 단계

- ◆ 학습된 지능에 기반하여 인지된 환경에 대한 추론 및 예측을 수행함
  - 훈련은 중앙의 데이터 센터에서 이루어짐
  - 추론은 대개 엣지(edge)의 스마트폰과 같은 디바이스에서 또는 니어 엣지(Near edge)에서 수행됨
- ◆ AI에서 프로덕션 배포는 모델로부터 인사이트를 얻을 수 있는 단계를 의미함
  - 이를 추론(Inference)이라고 함
  - 이 단계에서 딥러닝의 가치가 실현됨

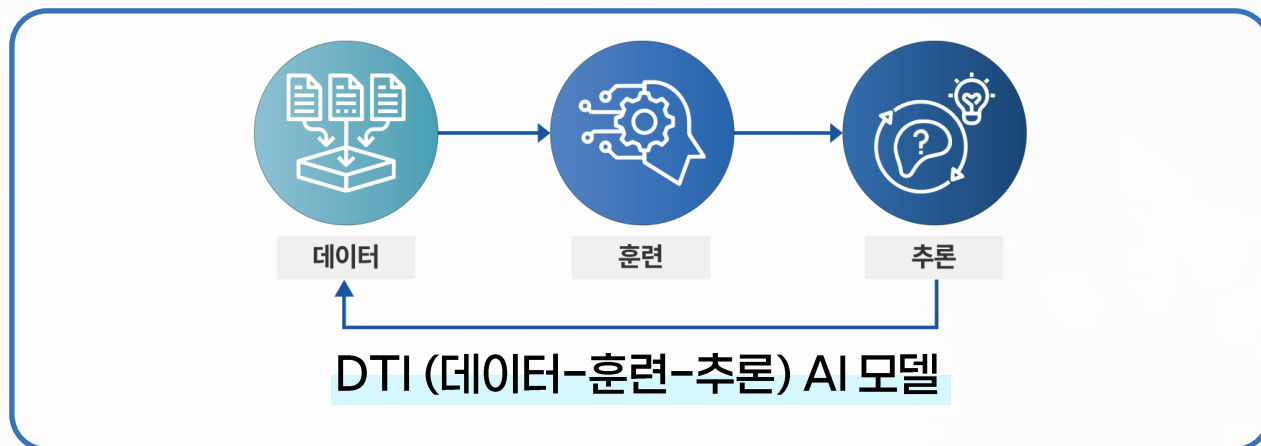


## 03 | 추론

### 추론 (Inference)

#### 3 추론 단계

- ◆ 추론 단계는 **모든 이전 요소들의 집합체임**
  - 데이터가 부족했거나 훈련이 부적절했다면 **추론**에 대한 **문제가 생김**
  - **추론**은 **대개 1초 이내**에 **끝나는 프로세스**이므로 빠르고 **정확한 인사이트**가 **필요함**
  - DTI 프레임워크에서 루프형태의 화살표가 보여주는 것처럼, 실제 환경에서 추론을 통해 **수집된 데이터**는 다시 **데이터 단계**의 워크플로우로 **피드백**됨







## 03 | 추론

### 추론 (Inference)

#### 3 추론 단계

- ◆ 데이터로부터 **학습된 신경망**은 **애플리케이션 형태**로 이미지, 음성, 혈액 질환 등을 인식해 내거나 누군가가 구입할 것 같은 제품을 제안하는 등 **다양한 영역**에서 **사용됨**
  - **훈련된 데이터**를 활용해 **구체적인 기능**을 **구현**하는 ‘**추론**’ 작업을 **해내는 것임**
  - 추론은 **소량**의 **데이터**만으로 무엇이 **정답**인지 정확하게 **예측**하는 것임
  - 음성으로 작동하는 스마트폰의 보조 기능부터 구글의 음성 인식, 이미지 검색, 스팸 필터링까지 **다양한 애플리케이션**에서 **추론**을 사용하고 있음
    - ▬ 바이두에서의 음성 인식, **페이스북**의 **이미지 인식**,  
아마존과 넷플릭스의 추천 엔진, **삼성전자**의 **스마트폰 S24** 등이 있음



## 03 | 추론

### 추론 (Inference)

△ 온 디바이스 스마트폰 “갤럭시 S24” 시리즈 공개

◆ ‘실시간 통화 통역’ 언어의 장벽 허무는 혁신

- 13개 언어 지원
- 문자와 주요 메신저에서 ‘메시지 번역’
- 카메라 촬영을 넘어 결과물 편집까지 AI가 사진 분석 후 맞춤형 편집 도구를 제안하는 ‘편집 제안’ 기능
- NPU 성능의 모바일 AP 기반 AI 프로세싱

