



<AI로 내년 연애 확률 예측하기> 세션에 참여해 주셔서 감사합니다. 세션 내용을 다시 정리해 보았습니다. 하단의 내용들을 통해, AI에 대해 더 자세히 이해하실 수 있으면 좋겠습니다.

1. AI란?

AI로 내년 연애 확률 예측해보기 앞서, AI가 무엇인지 먼저 알아 봅시다.



AI(인간 지능)란 무엇일까요?

AI를 직역하면 인간 지능입니다. 정말 말 그대로, 인간의 지능처럼 작동하는 것. 사람처럼 생각하는 것을 의미합니다.

2. AI 제작 과정

사람처럼 생각하는 것, AI를 만들기 위해 무엇을 해야 할까요?

사람이 생각할 수 있게 되는 과정과 유사합니다.

인간이 고양이와 개를 판단할 수 있게 되는 과정은 다음과 같습니다.

고양이나 개를 처음 접했을 때, 이게 '고양이다.' '개다.' 같이 바로 무엇인지 인지하지 못합니다.

하지만 고양이와 개를 많이 접하고, 부모님이나 선생님께 고양이와 개가 무엇인지 배우게 됩니다.

이렇게 많은 고양이와 개 예시들을 접하며 고양이와 개에 대해 학습하게 됩니다.

이 학습한 내용을 바탕으로 처음보는 외모의 고양이를 보아도, 처음보는 외모의 개를 봐도 '아! 재는 고양이네.' 와 같이 판단할 수 있게 됩니다.

AI 제작 과정도 유사합니다.



개와 고양이인지 판단할 수 있는 AI 제작 과정을 예로 들어 보겠습니다.

개와 고양이 사진, 개와 고양이의 특징이 담긴 글 등 개/고양이에 대한 정보(데이터)가 제공되어야 합니다.

이런 정보들을 바탕으로 '개가 무엇인지', '고양이가 무엇인지' 학습하게 됩니다.

학습한 걸 바탕으로 새로운 고양이 데이터를 주어도, 고양이인지 잘 판단하는지 평가하는 과정을 거치게 됩니다.

평가 과정에서 좋은 성과가 확인되면, 해당 AI를 배포해 새로운 고양이/개 사진들을 판단하는 데 사용합니다.

Cf)

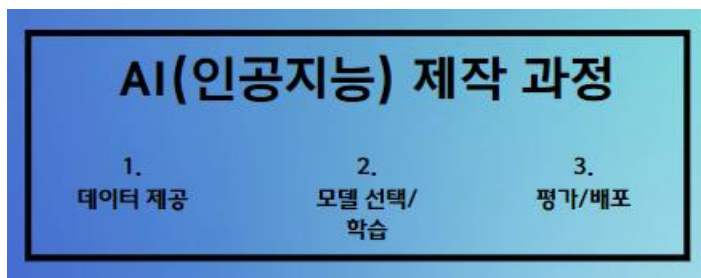
인간의 학습 과정	AI 제작 과정
개와 고양이 접하기	개와 고양이 데이터 제공받기

개와 고양이가 무엇인지 배우기. (개와 고양이의 특징을 파악.)	개와 고양이가 무엇인지 학습. (개와 고양이의 특징을 학습 / AI모델이 특징을 추출해 냄.)
새로운 개와 고양이를 봐도 무엇이 개인지 고양이인지 판단 가능.	새로운 개와 고양이 데이터가 제공돼도 무엇이 개인지 고양이인지 잘 파악하면 -> 여러 사람들이 해당 AI를 사용해 개와 고양이를 판단할 수 있도록 서비스 배포

3. '연애 예측 AI 만들기'

AI 제작 과정을 알아 보았습니다.

이 과정에 따라 '연애 확률 예측 AI'를 만들어 볼 수 있습니다.



연애 예측에 필요한 데이터를 만든 후, AI에 데이터를 제공해 봅시다. 그 다음, 데이터를 학습시키고, 배포 과정까지 거쳐 봅시다.

어떻게 데이터를 만들고, 학습시켜, 평가/배포까지 할 지 막막할 것입니다.

메모장으로 빠르게 데이터를 만든 후, Azure AI라는 Microsoft의 서비스를 통해 2&3번 과정을 쉽게 수행할 수 있습니다.

[실습.pdf] 에 따라 '연애 확률 예측 AI'를 만들어 보세요.

4. Azure AI – machine learning studio

실습을 위해 Azure AI 서비스를 이용했습니다. Azure AI 서비스 중에서도 Machine Learning Studio 서비스를 사용했습니다.

Machine Learning Studio에서 실습에서 만든 숫자값 예측뿐만 아니라, 이미지를 처리할 수 있는

AI, 언어를 처리할 수 있는 AI 등등의 여러 기능들을 제공합니다.

Machine Learning Studio 서비스들을 하단의 링크에서 무료로 배우실 수 있습니다.

<https://learn.microsoft.com/ko-kr/azure/machine-learning/overview-what-is-azure-machine-learning?view=azureml-api-2>

하단의 링크를 통해 Azure Machine Learning Studio 서비스를 배운 후, 여러 AI들을 직접 만들어 보길 바랍니다.

5. AI의 다양한 분류 기준

chatGPT를 AI 모델 gpt 이렇게 부르기도 하지만, 멀티모달 모델 chatGPT, 딥러닝 모델 ChatGPT, 강화학습 모델 ChatGPT 등 정말 다양한 방법으로 ChatGPT를 부릅니다.

이렇게 다양하게 AI를 부르는 기준이 있습니다. 경북대학교를 위치에 따라서는 지방대학교, 설립 주체에 따라서 국립대학교와 같이 다양하게 분류하고 부를 수 있습니다.

이처럼 AI 또한 어떤 식으로 분류하냐에 따라 다양하게 불릴 수 있습니다.

이 문서에선 총 3가지 분류 기준을 설명 드리고자 합니다.

(1) '사람의 개입 여부'

인공지능 관련해 딥러닝, 머신러닝이라는 용어를 많이 들어 보셨을 겁니다. 이 용어는 다음과 같은 포함관계를 가지고 있고, '사람의 개입 여부'에 따라 더욱 특정한 AI 범위를 칭하고 있는 것입니다.

딥러닝 ⊂ 머신러닝 ⊂ 인공지능

머신러닝은 컴퓨터에게 데이터를 보여주고, 컴퓨터가 그 데이터로부터 규칙을 스스로 찾아내도록 하는 것입니다. 사람은 규칙을 직접 알려주지 않아도 되고, 컴퓨터가 데이터를 보고 학습합니다.

예를 들어, 사진에서 고양이와 강아지를 구별하게 하는 것도 머신러닝의 일입니다.

딥러닝은 머신러닝 중의 하나로, 인공신경망을 사용하는 것입니다.

인공신경망은 컴퓨터의 뇌처럼 동작하는 모델입니다.

머신러닝에서는 사람이 AI 모델의 구조와 특징을 직접 정해주어야 하지만, 딥러닝에서는 모델이 스스로 특징을 학습하므로 사람의 개입이 상대적으로 적다는 게 차이점입니다.

(2) '정보 제공 정도'

2. '학습 데이터 정보 제공 정도'에 따른 AI 분류



'좌측은 고양이, 우측은 개'
데이터 정보 제공



정보 제공 X
알아서 특징들을 추출.



보상 시스템

정보 제공 정도에 따라 AI를 지도학습 AI/ 비지도 학습 AI/ 강화 학습 AI로 분류할 수 있습니다.

예를 들어, 설명 드리겠습니다.

개와 고양이 사진이 있을 때, 좌측은 고양이고 우측은 개라고 정보를 제공해주며 지도해주는 것이 지도학습입니다.

개와 고양이 사진을 주고, 알아서 개와 고양이가 무엇인지 스스로 특징을 분석해, 학습하라는 것이 비지도 학습입니다.

강화 학습은 개와 고양이 사진을 먼저 주고 일단 무엇이 고양이인지 맞춰 보라고 하는 것입니다. 좌측 사진이 고양이라며 정답을 말하면 가산점을 줍니다. 이와 같이 보상을 주며 올바른 답을 할 수 있는 AI를 만드는 방법이 강화 학습입니다.

(3) '학습 데이터 종류'

3. '학습 데이터 종류'에 따른 AI 분류



CV(Computer Vision)

이미지/ 영상 등 시각적 정보



LM(Language Model)

언어 모델



Multi Modal

다양한 데이터 종류

다양한 학습 데이터 종류에 따라 인공지능을 다양하게 부를 수 있습니다.

1) 컴퓨터 비전 (Computer Vision - CV) : 시각 데이터 처리 AI 모델

컴퓨터 비전은 컴퓨터가 이미지나 비디오를 보고 이해하는 것을 말합니다.

예를 들어, 얼굴을 인식하거나 고양이와 강아지를 구별하는 것처럼 시각적 정보를 처리하는 기술입니다.

인공지능이 시각 정보를 다루면서 "시각 인공지능" 또는 "컴퓨터 비전 AI"라고 부릅니다.

2) 언어 모델 (Language Model - LM) : 언어 처리 AI 모델

언어 모델은 컴퓨터가 언어를 이해하고 생성하는 데 사용되는 특별한 유형의 인공지능 알고리즘 또는 모델입니다.

언어 모델은 주로 텍스트 데이터를 기반으로 학습되며, 문장 또는 문서의 문맥을 파악하고 다음 단어를 예측합니다.

예를 들어, "나는 학교에 갔다"라는 문장에서 "학교에" 다음에 올 단어로 "컴퓨터"를 예측할 수 있는 것이 언어 모델의 역할입니다.

3) 멀티모달 (Multi-Modal) : 혼합된 데이터 처리 AI 모델

멀티모달은 여러 종류의 데이터를 함께 사용하여 인공지능을 훈련하고 작업을 수행하는 것을 의미합니다. 이 데이터는 이미지, 텍스트, 음성 등 다양한 형태일 수 있습니다.

예를 들어, 스마트폰에서 음성으로 질문하면 화면에 정보를 표시하는 것은 멀티모달 AI의 한 예입니다.

6. 주식/뉴스의 핫한 AI 용어

AI가 무엇인지 앞의 내용들을 통해 이해가 되셨나요? AI가 무엇인지 이해하셨길 바라며, 마지막 내용으로 AI 동향을 다뤄 보려 합니다.

요즘 주식/뉴스에서 핫한 AI 단어들을 설명드리며 가볍게 동향 또한 함께 설명드리고자 합니다.

주식/뉴스 핫 토픽 AI 용어



세션 전반부 용어

LM, CV, 멀티 모달 등



생성 AI(GAI)

Generative AI

예 : chatGPT, DALL-E



LLM vs sLLM

Large Language Model/
small Large Language Model

범용적 vs 맞춤형

1) 멀티 모달

앞 전에서 LM, CV, 멀티 모달에 대해 설명드렸습니다. 설명드렸다고 멀티 모달은 다양한 데이터를 처리하는 AI 모델입니다. 언어에 대해서 처리하는 LM, 시각적인 데이터에 대해서 처리하는 CV보다 다양한 데이터를 함께 처리할 수 있는 멀티 모달 AI를 만드는 게 더 어렵겠죠?

그런 만큼 요즘 기업들은 멀티 모달 모델을 출시하고 있고, 출시하기 위해 노력 중입니다.

2) 생성형 AI

예전의 챗봇은 '오늘 밥 뭐 먹어'를 계속 입력했을 때 '개구리반찬' 등 과 똑 같은 답만 계속 나왔습니다. 하지만, chatGPT는 '오늘 밥 뭐 먹어'를 입력했을 때마다 항상 똑 같은 답을 주진 않습니다. 질문에 대한 답을 생성해낼 수 있는 이러한 AI를 생성형 AI라고 부릅니다.

인간만의 영역이라고 생각했던, 생성해낼 수 있는, 생각해 낼 수 있는 능력을 가졌다는 점에서 생성형 AI는 AI 등장때부터 현재까지 계속 큰 화젯거리입니다.

3) LLM / sLLM

앞 전에 언어 모델(LM)에 대해 설명드렸습니다. LM은 언어를 처리하는 AI를 칭하는 것이고, LLM은 Large Language Model, 대규모 언어 모델을 칭합니다.

방대한 양의 텍스트 데이터를 학습하여 자연어를 생성하고, 언어를 번역하고, 다양한 종류의 창의적인 콘텐츠를 작성하고, 질문에 답변하는 등 다양한 작업을 수행할 수 있는 AI 모델입니다.

sLLM은 Small Large Language Model의 약자로, LLM에 비해 작동할 때 판단의 바탕이 되는 것, 학습에 필요한 데이터 수가 적은 소형 언어 모델입니다.

즉, LLM은 sLLM보다 더 많은 데이터를 학습하고, 더 복잡한 작업을 수행할 수 있지만, 학습 시간이 오래 걸리고, 컴퓨팅 자원(연산할 수 있는 것)이 많이 필요합니다. 반면, sLLM은 LLM에 비해 학습 시간이 짧고, 컴퓨팅 자원이 적게 필요하지만, LLM에 비해 성능이 떨어질 수 있습니다.

LLM은 보통 chatGPT와 같은 범용적인 서비스가 그 예입니다. sLLM은 LG 등의 기업에서 회사 내부 직원들만 사용하는 챗봇 등 한정된 상황에서 서비스하기 위해 많이 사용됩니다.

LLM을 바탕으로 sLLM을 만들거나, 회사에 알맞은 sLLM을 만들어 주는 특정 회사들이 요즘 뉴스와 주식에서 화제가 되고 있습니다.

LLM을 만들고 서비스를 제공하는데, 정말 많은 데이터와 인력이 필요하기 때문에 대기업을 제외하곤 LLM 서비스를 만들기조차 쉽지 않습니다.

그래서 상대적으로 다양한 대상들이 서비스를 만들어 낼 수 있는 sLLM에 많은 관심이 쏠리고 있습니다.