



Introducing Agent AI on scholar system in UNIST

August 30, 2024

Department of Mathematical science

20241341 Chiwoo Jeon

CONTACT

Ulsan National Institute of Science and Technology

Address 50 UNIST-gil, Ulju-gun, Ulsan, 44919, Korea

Tel. +82 52 217 0114 **Web.** www.unist.ac.kr

Mathematical science

Building 110 Room U207

목차



1. Intro

2. 최근 인공지능 업계 동향

3. Coredottoday X UNIST - Math tutor AI



Intro

FIRST IN CHANGE

Introduce our team(수리과학부 김필원 교수 소모임)

1. Model development

- 교내 학술자료 한-영 기계번역 모델
- 교내 학습 도우미 자기회귀 모델

2. AI Deployment

- Math AI Tutor by using RAG Agent System

Introduce our team



응용수학 소모임

MAIN PROCESS

MODEL DEVELOPING	TRAIN PIPELINE (BASELINE)	INFERENCE PIPELINE	DATA PRE-PROCESSING	PRE-TRAINING	FINE-TRAINING
학술 한-영 번역 모델 (Encoder-Decoder Transformer)	✓	✓	✓	✓	✗
자기회귀모델(Decoder-only Transformer, GPT)	✓	✓	✓	✗	✗
SYSTEM AI	PIPELINE	BACKEND	FRONTEND	DEPLOYMENT	
Math AI Tutor	✓	✗	✗	✗	
교내 인공지능 시스템					

Introduce our team

1. Baseline model of neural machine translation

Model Architecture	Transformer
d_model	512
Decoder_num, Encoder_num	6, 6
Tokenizer	Mecab + BPE
Optimzer	AdamW
Activate_function	GELU(Gaussian Error Linear Unit)
Head_num	8
Total_parameter_num	≈77840000

```
bleu_test.py > ...
93     use_GELU=config.use_GELU,
94     max_length=config.block_size,
95 )
96
97 model_2 = Transformer(
98     input_size=config.input_size,
99     d_model=config.d_model,
100    output_size=config.output_size,
101    head_num=config.head_num,
102    encoder_num=config.n_encoder,
103    decoder_num=config.n_decoder,
104    dropout_p=config.dropout_p,
105    use_GELU=config.use_GELU,
106    max_length=config.block_size,
107 )
```

```
(GPT) (base) letsgotothehanriver@chiwoonvidia:~/transformer$
```

Introduce our team

src: Considering the perspective of the company's risk management, AAA 1 divided the risk classification; strategic risks, business risks, capital and accounting risks, infrastructure risks and safety, and health risks, and AAA 2 stated that specific strategies such as improving awareness, improving the employee experience and expertise, promoting risk management services, and risk management are needed to manage risks for companies.666

tgt: 기업의 리스크 관리에 대한 선행연구의 견해를 살펴보면, AAA 1은 리스크의 구분을전략 리스크, 사업 리스크, 자본·회계 리스크, 인프라 리스크, 자연재해 리스크, 안전·건강리스크 등으로 구분하였고, AAA 2는 기업의 리스크 관리를 위해서는 인식개선, 종사자의 경험과 전문성 개선, 리스크 관리에 대한 정보교류 활성화, 리스크 관리 서비스 강화,리스크 관리 교육제도와 같은 구체적인 전략이 필요함을 밝혔다.

pred: AAA 1은 기업의 리스크 관리 관점을 고려하여 위험분류, 사업 리스크, 자본 및 회계 리스크, 인프라 위해, 건강 리스크, 인프라 위험, 건강 리스크의 구분을, AAA 2는 기업의 리스크관리를 위해서는 인지 개선, 직원의 경험과 전문성 제고, 리스크관리 서비스 육성, 리스크관리 서비스, 리스크 관리 등의 구체적인 전략이 필요함을 강조한다고 하였다.

Introduce our team

src: Rights to education on legal-oriented areas of basic youth labor rights, rights to the detailed environment, rights to education on social security systems, especially rights to labor rights, have been established, but relevant laws are insufficient, inappropriate ways.

tgt: 법령 중심의 청소년 노동기본권 영역과 세부내용 환경에 대한 권리 정당한 사회보장 체계에 대한 권리 대해 교육받을 권리 특히, 노동기본권에 대해 교육 받을 권리는 아동권리협약과 청소년기본법 등 관련 법령이 마련되었으나, 적절한 방식으로 교육 받을 권리에 있어서는 관련 법령이 미비한 것을 알 수 있다.

pred: 청년 노동 기본권의 법 중심의 교육에 대한 권리, 상세환경, 사회보장제도, 특히 노동 권리에 대한 교육에 대한 권리를 규정하고 있지만, 관련 법률이 미비해 부적절한 방법이다.

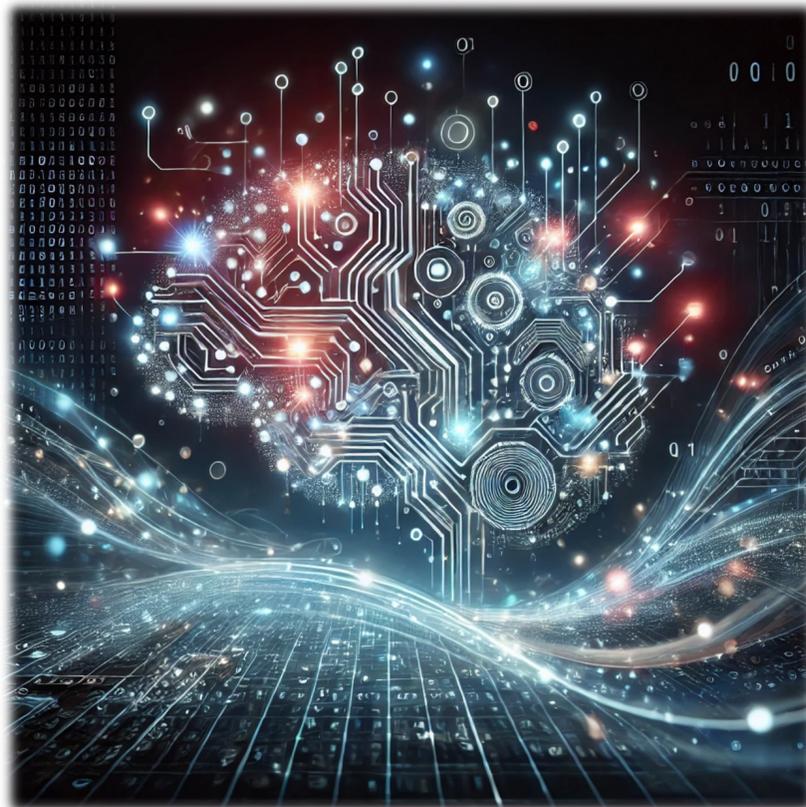


About Trend

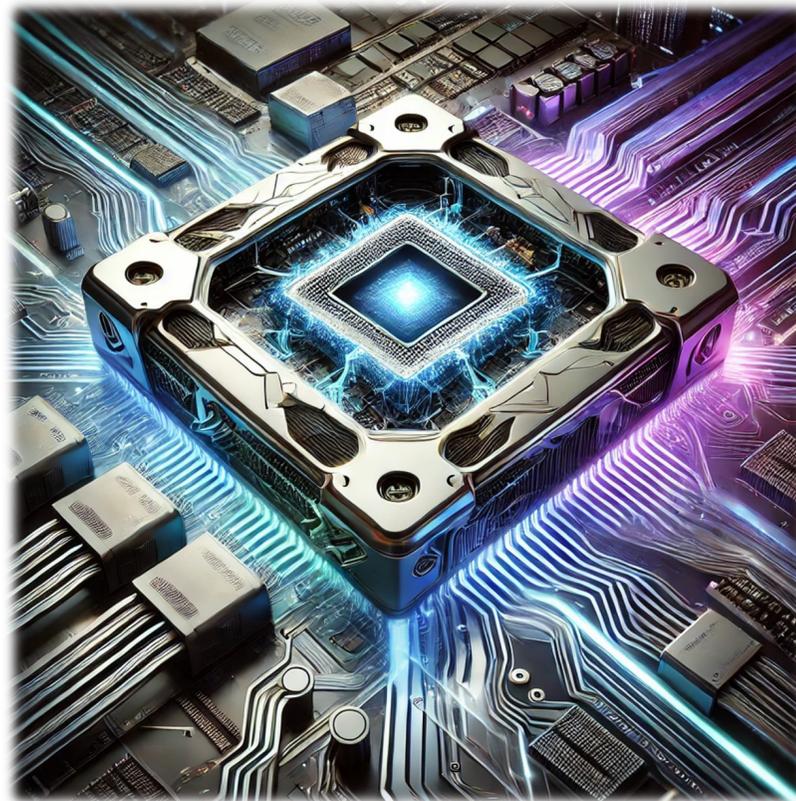
FIRST IN CHANGE

Main 3 Component in developing AI

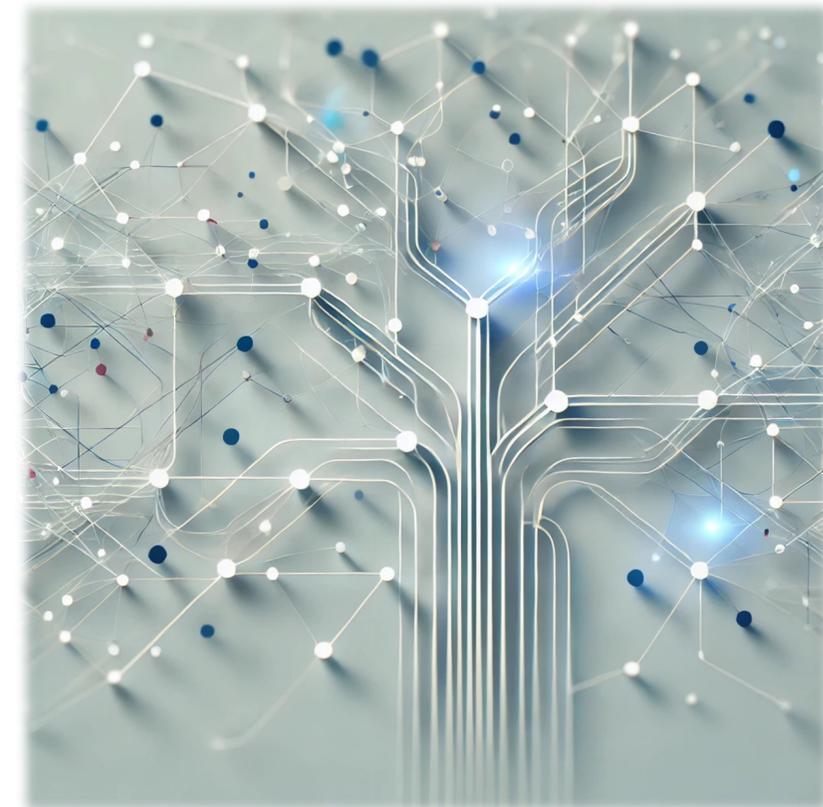
BIG DATA



GPU

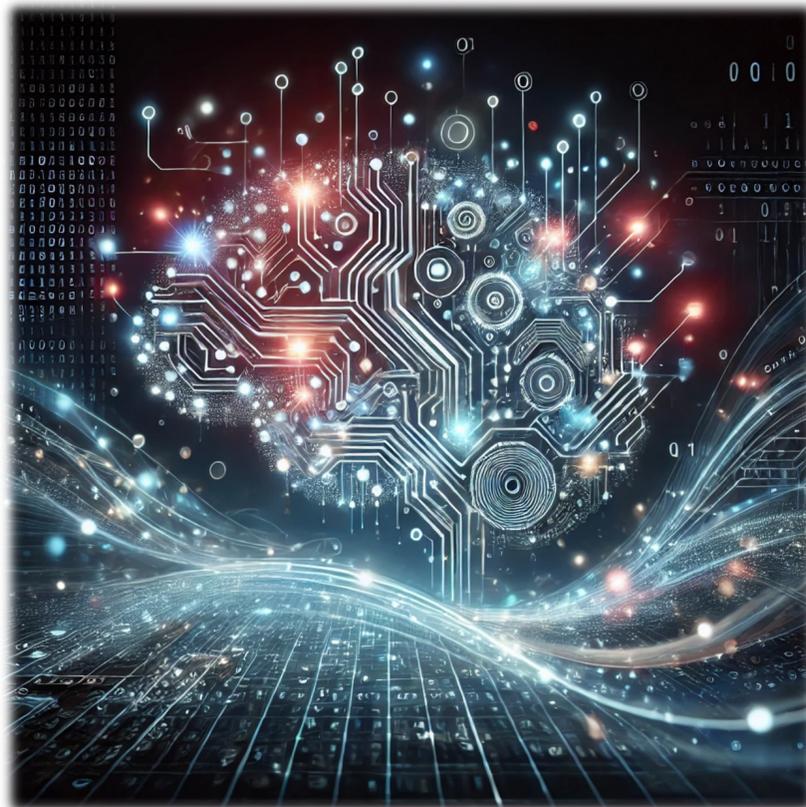


Algorithm

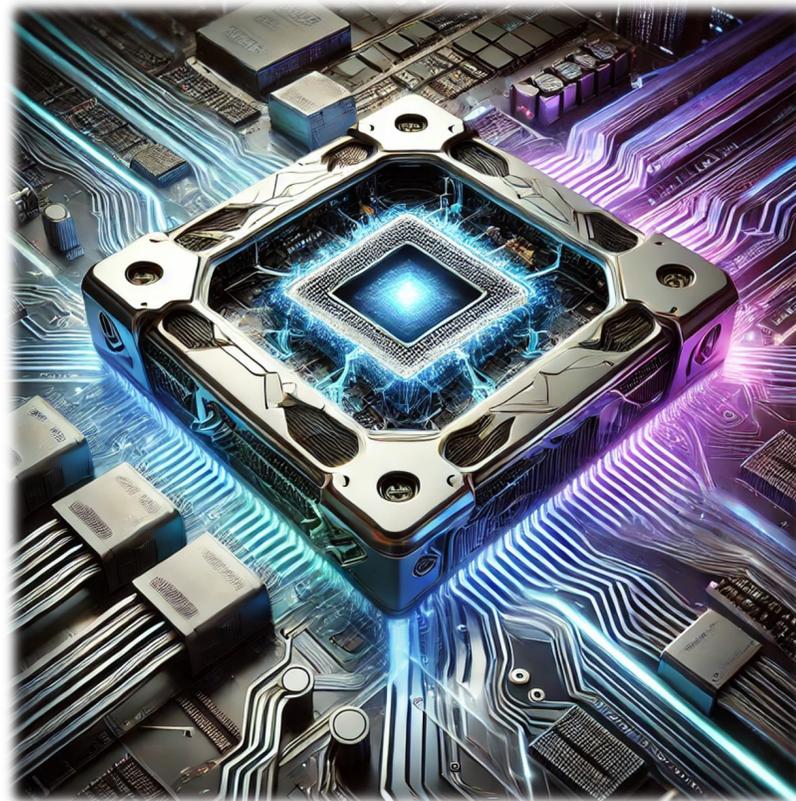


Main 3 Component in developing AI

BIG DATA



GPU



Algorithm

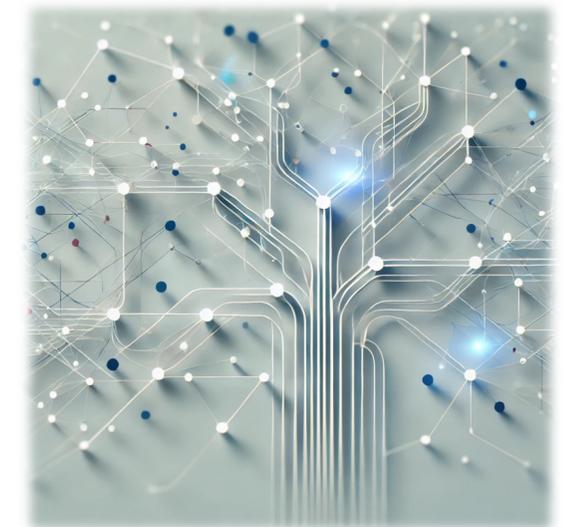
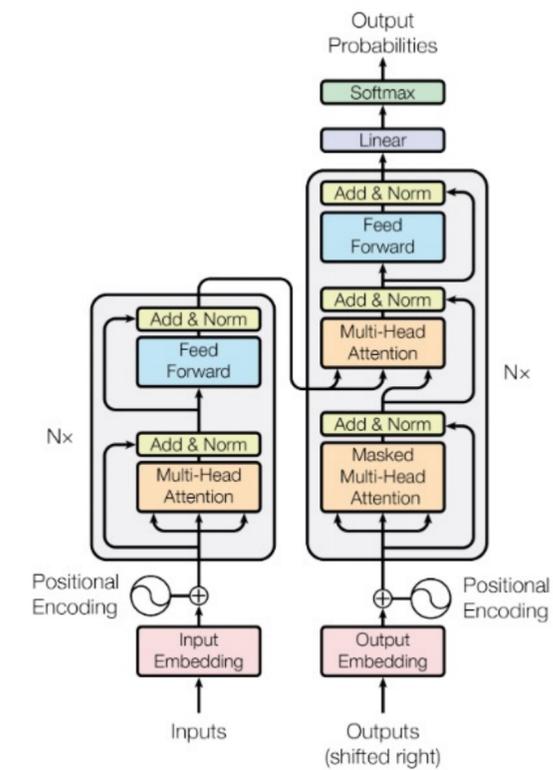


Figure 1: The Transformer - model architecture.

Main 3 Component in developing AI

BIG DATA



kakao

Meta

Microsoft

GPU



Algorithm

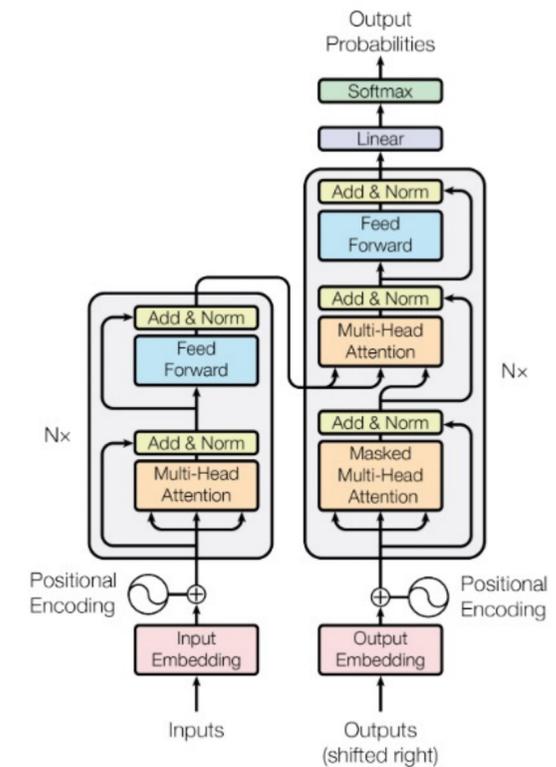


Figure 1: The Transformer - model architecture.

Main 3 Component in developing AI

BIG DATA



GPU



Algorithm

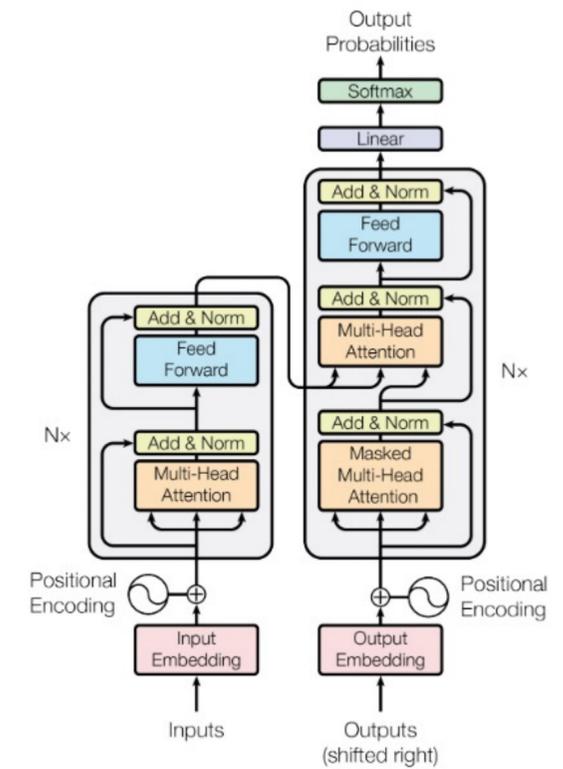


Figure 1: The Transformer - model architecture.

Main 3 Component in developing AI



BIG DATA

GPU

Algorithm

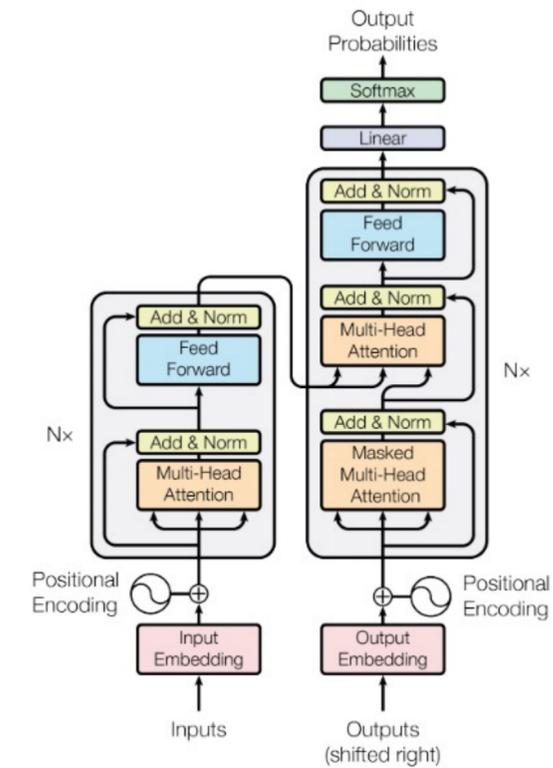
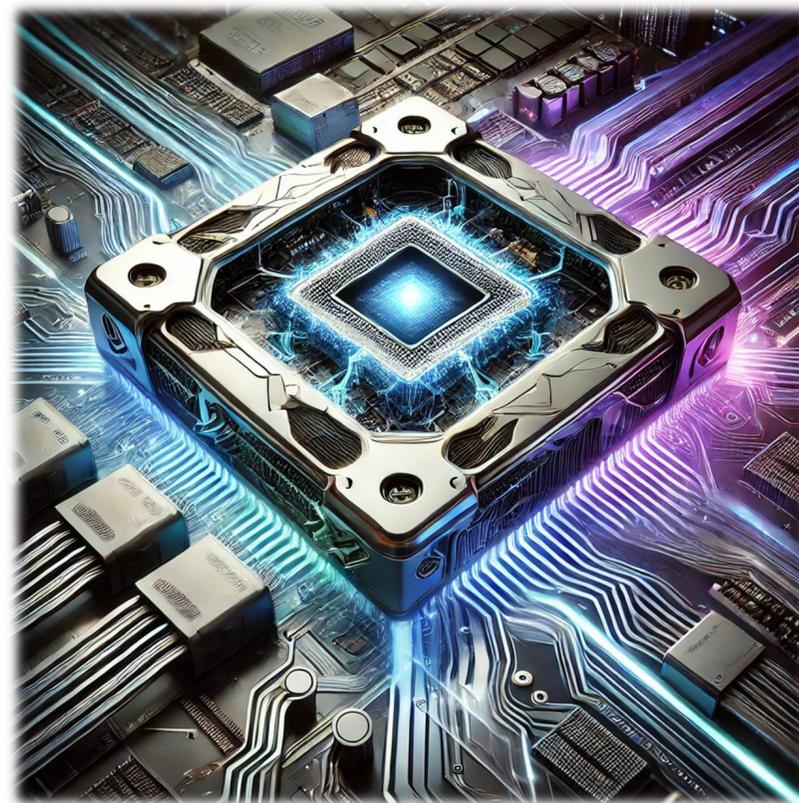
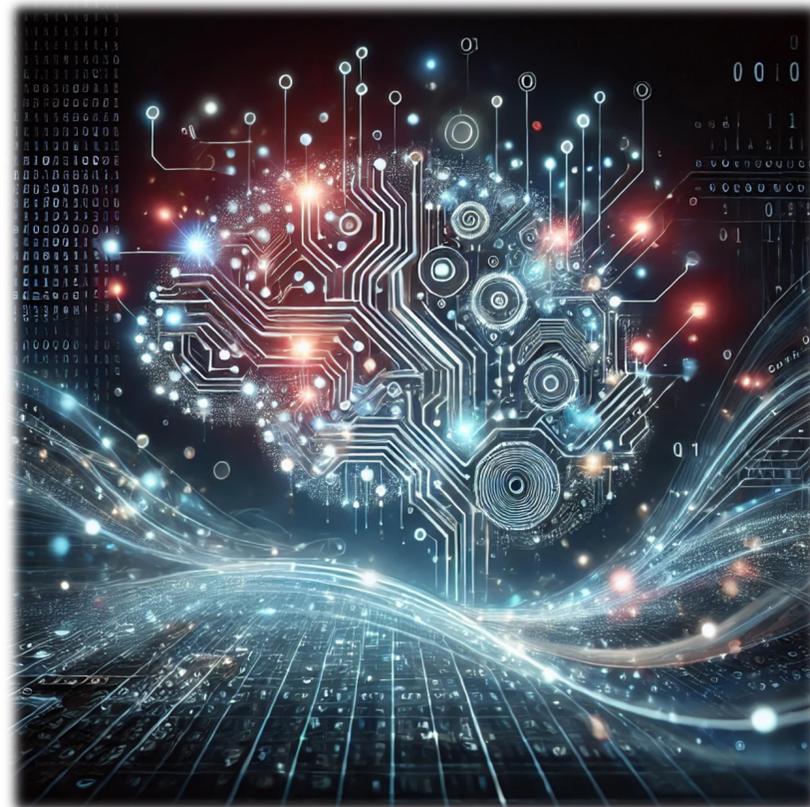
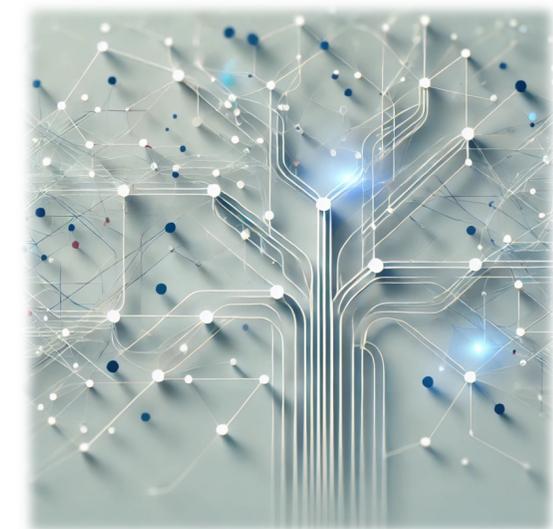


Figure 1: The Transformer - model architecture.

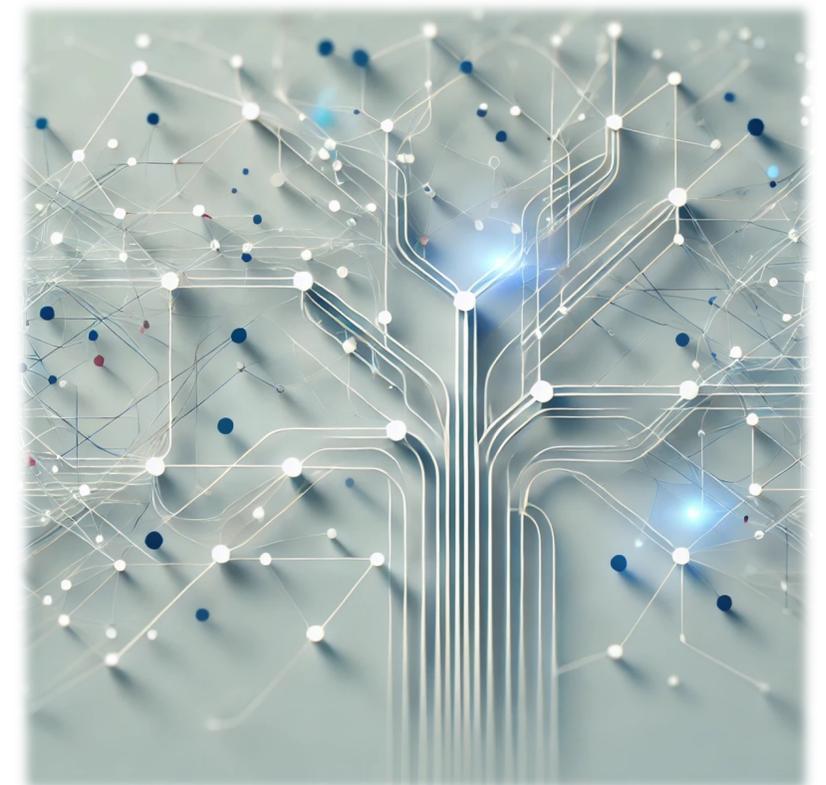


Main 3 Component in developing AI

 Gemini

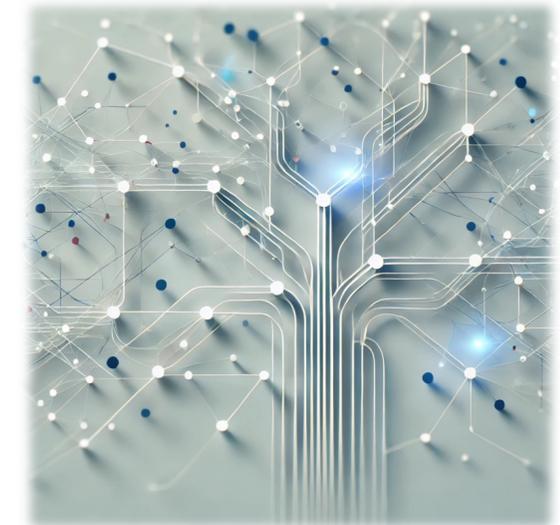


Algorithm



Main 3 Component in developing AI

Algorithm

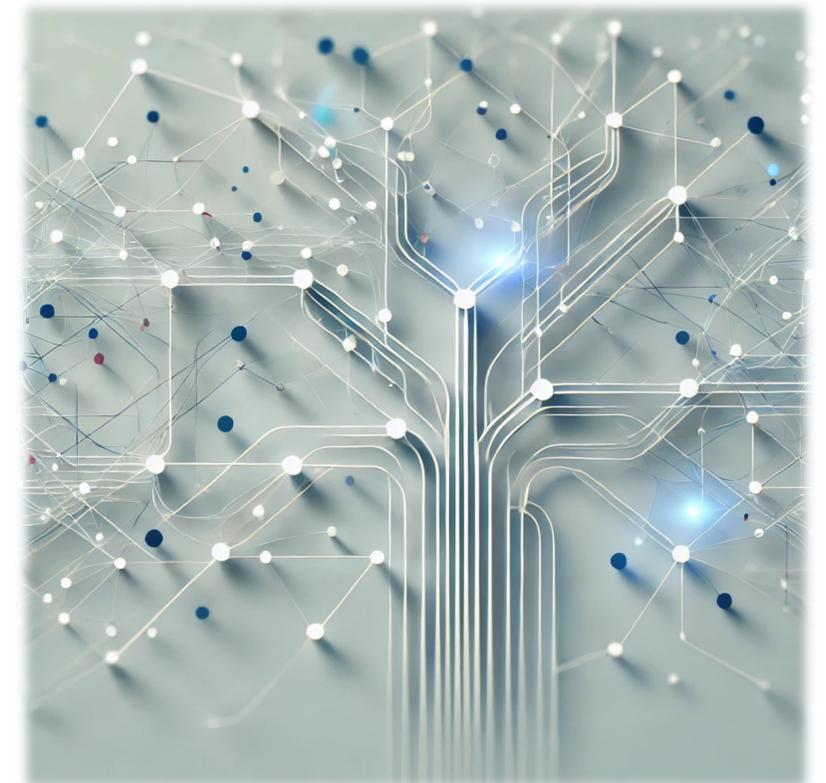


Main 3 Component in developing AI

 Meta

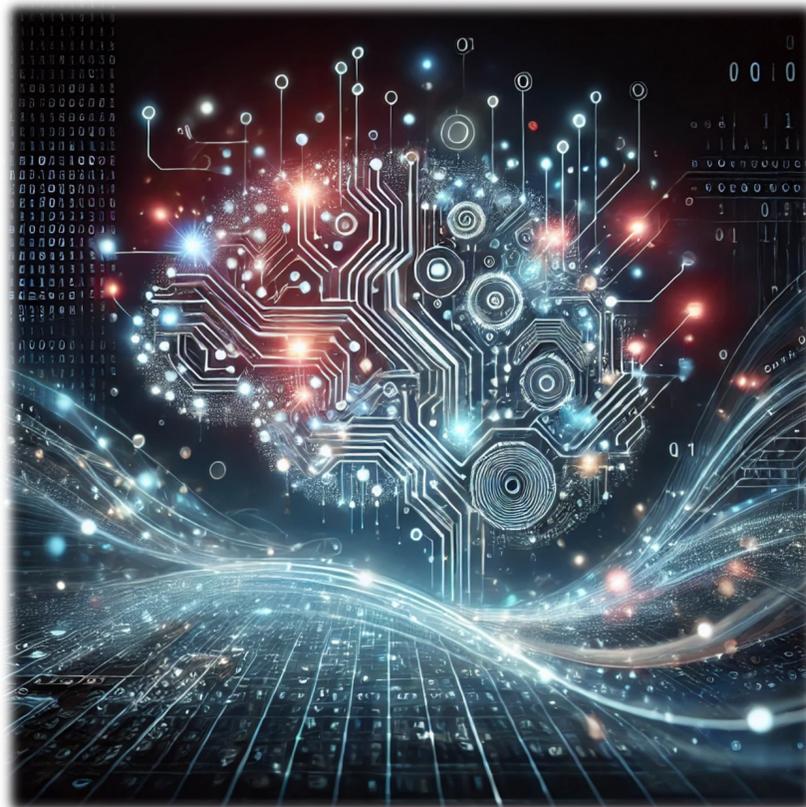


Algorithm

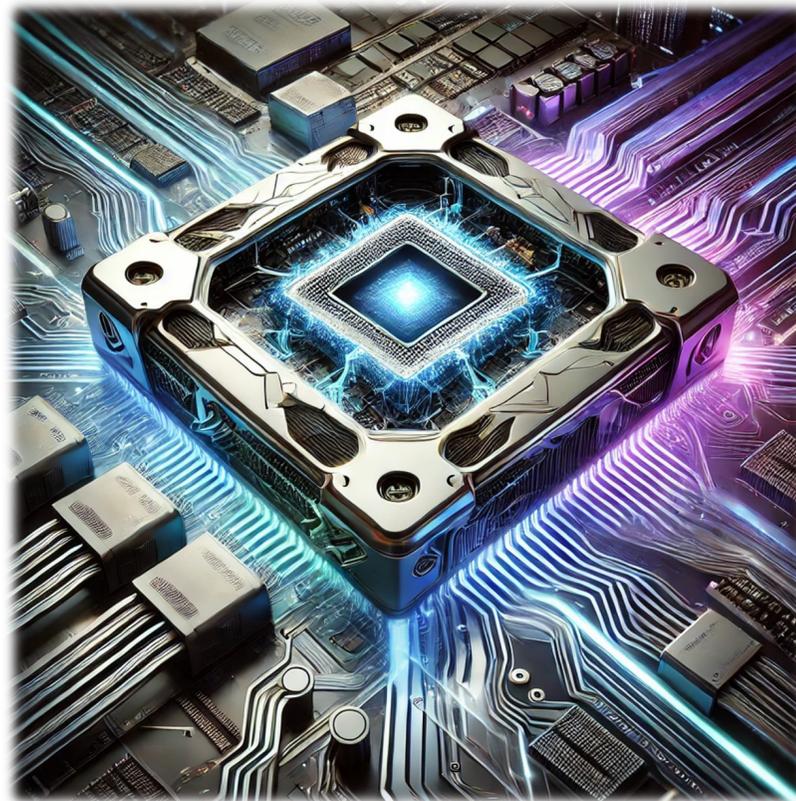


Main 3 Component in developing AI

BIG DATA



GPU



Algorithm

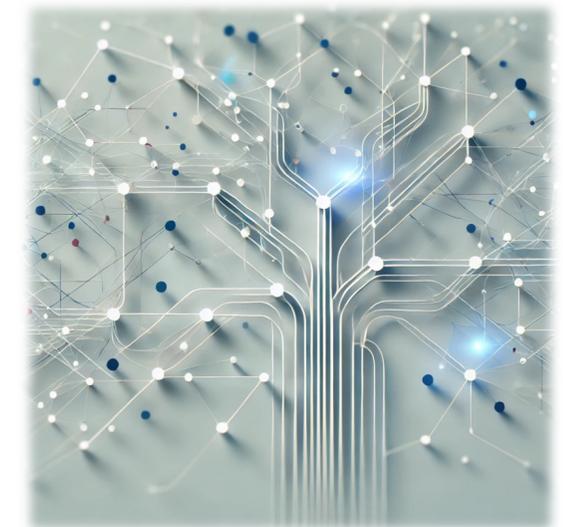
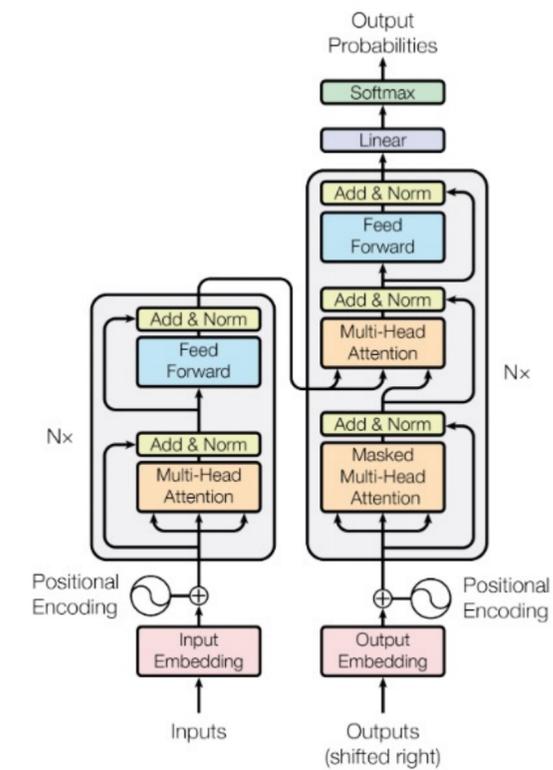


Figure 1: The Transformer - model architecture.

Main 3 Component in developing AI



BIG DATA

GPU

Algorithm

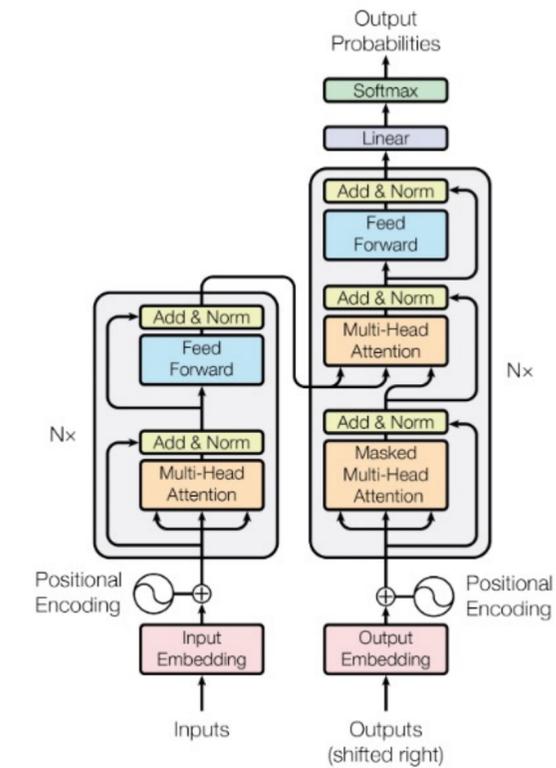
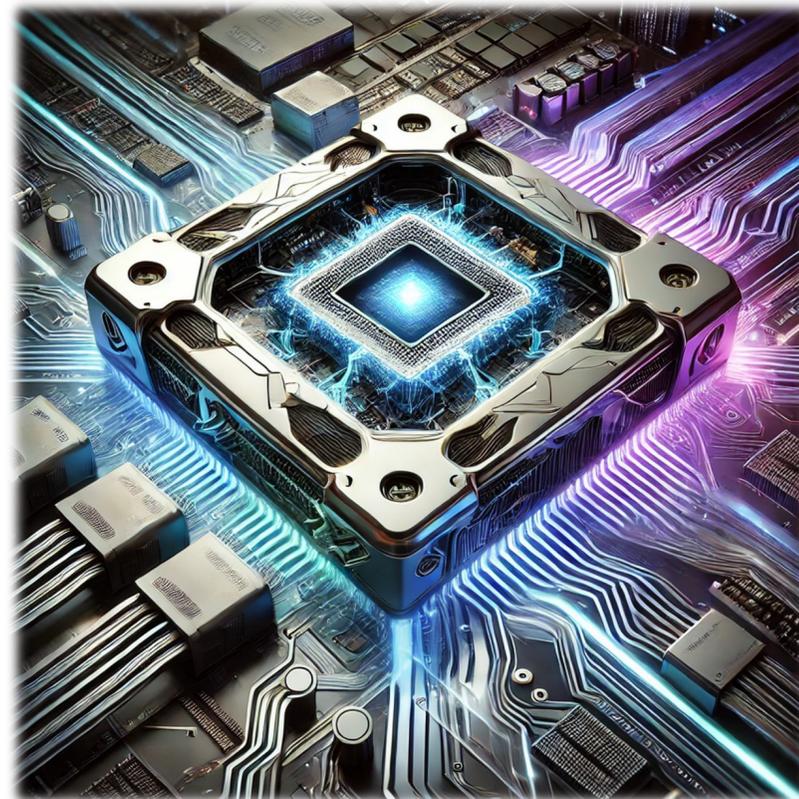
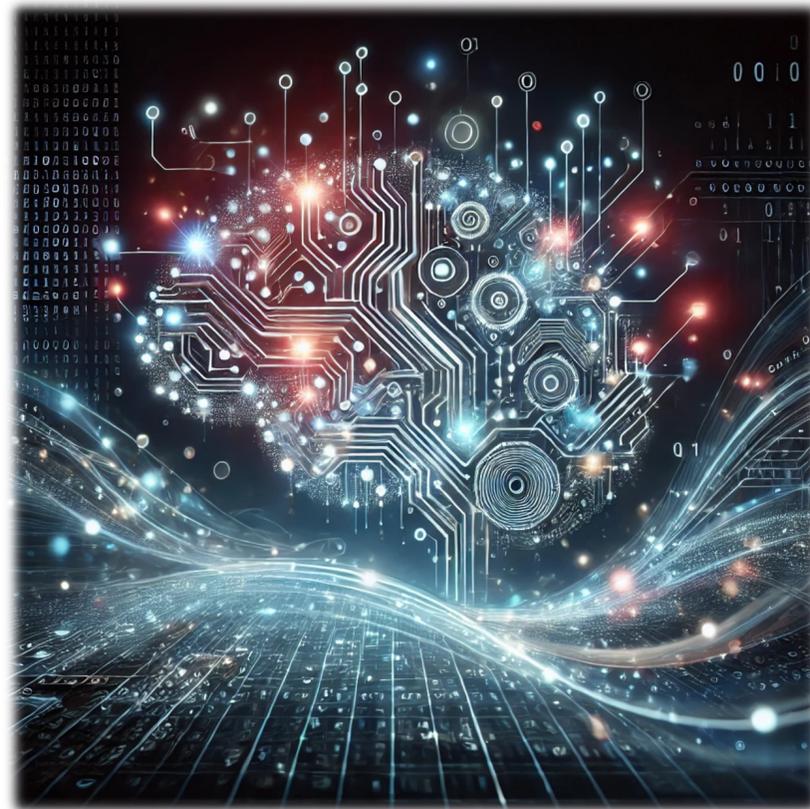
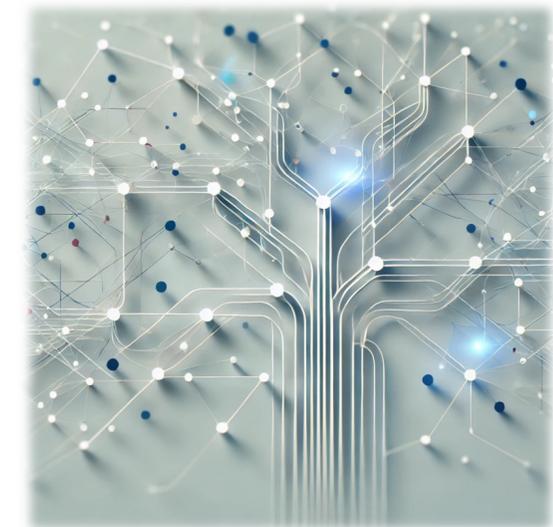


Figure 1: The Transformer - model architecture.





Core-dot today
X
UNIST Math

FIRST IN CHANGE

Mathematical Approach

현재의 자연어처리 모델(GPT)은 인공 “지능” 보다는 단일 함수에 가깝다

한 마디로 자연어생성 모델을 정의하자면?

“T-1 시점까지의 토큰을 조건부로, 맥락적으로 가장 올바른 T시점 토큰의 확률 분포를 예측하는 고차원 다변수함수”

“지능”이란, 여러 함수의 집합체이자 단일 함수

$$\hat{y}_t = \arg \max_{y_t} P(y_t | \mathbf{y}_{1:t-1}) \quad P(\mathbf{y}) = \prod_{t=1}^T P(y_t | y_1, y_2, \dots, y_{t-1})$$

Mathematical Approach

입력: 유니스트에 대해서 설명해줘

출력:

- Step 1: **유니스트** (조건: 유니스트에 대해서 설명해줘)
- Step 2: **는** (조건: 유니스트에 대해서 설명해줘, 유니스트)
- Step 3: **대한** (조건: 유니스트에 대해서 설명해줘, 유니스트는)
- Step 4: **민국** (조건: 유니스트에 대해서 설명해줘, 유니스트는 대한)
- Step 5: **의** (조건: 유니스트에 대해서 설명해줘, 유니스트는 대한민국)
- Step 6: **과학** (조건: 유니스트에 대해서 설명해줘, 유니스트는 대한민국의)
- Step 7: **기술원** (조건: 유니스트에 대해서 설명해줘, 유니스트는 대한민국의 과학)

Mathematical Approach

단순 인공지능 자연어 생성 모델은 수학적 함수처럼 입력에 대해 고정된 출력을 제공

> 훈련된 데이터에만 의존하여 새로운 입력에 대해 적절하지 않은 답변을 할 수 있음

이를 보완하기 위해 RAG를 활용한 Agent 시스템을 구현할 수 있음

> 외부 지식베이스나 실시간 데이터를 이용해 입력에 대한 보다 정확하고 풍부한 출력을 제공

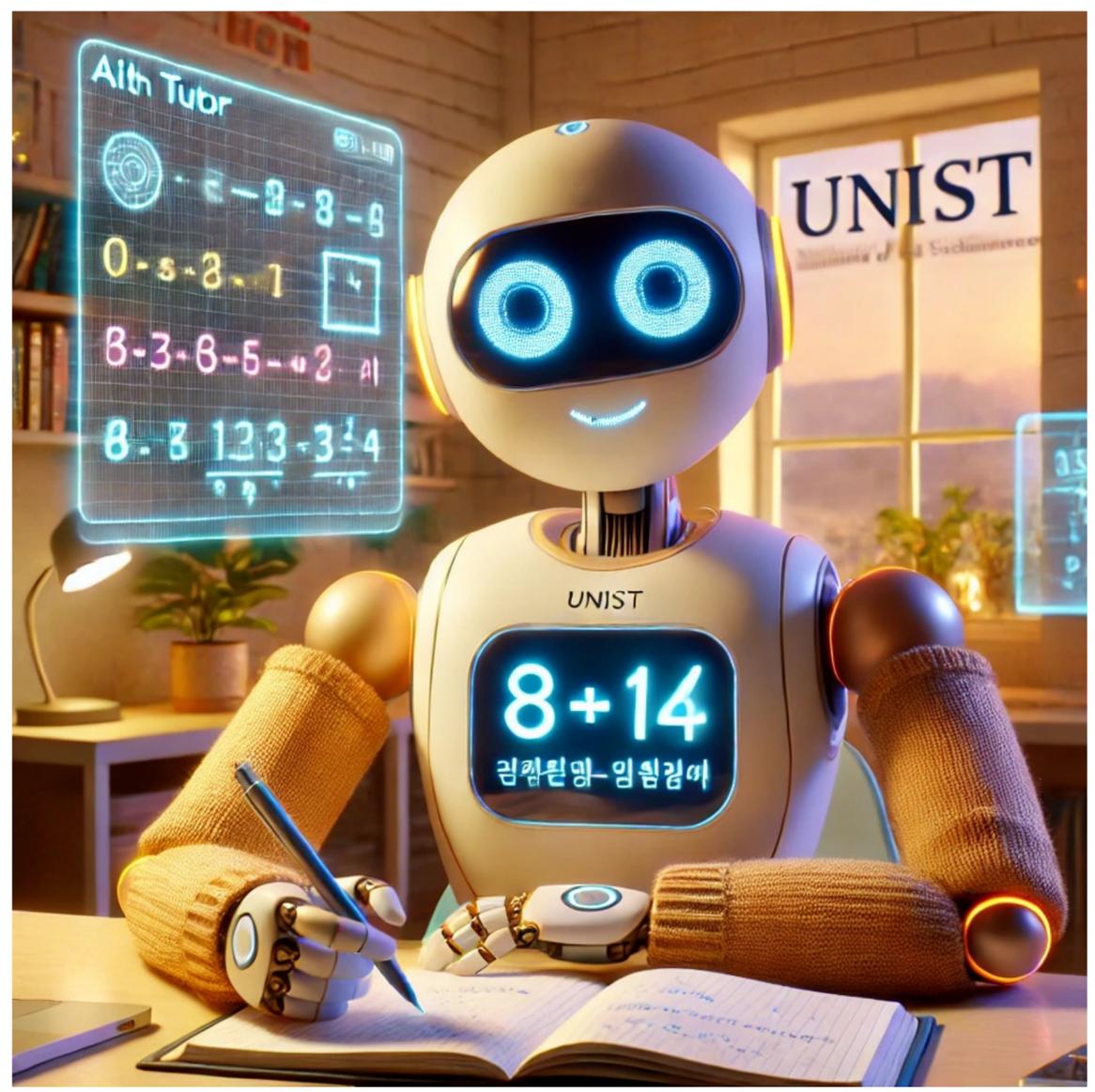
➤ 내부적 회귀 알고리즘을 통해 최적의 문제해결 결과를 도출할 때 까지 실행을 반복할 수 있음

➤ AGI의 핵심 알고리즘으로 예상

➤ **Function (x) Agent(O)**

Mathematical Approach

앞선 섹션에서 소개한 인공지능 시스템 개발의 편의성 제고에도 불구하고, 현재 국내 국가 및 대학 등 공공기관에서 내부 Agent AI를 도입하려는 시도는 전무함



UNIST

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

CORE.TODAY

문제 정의:

> 훈련된 데이터에만 의존하여 새로운 입력에 대해 적절하지 않은 답변을 할 수 있음

이를 보완하기 위해 RAG를 활용한 Agent 시스템을 구현할 수 있음

> 외부 지식베이스나 실시간 데이터를 이용해 입력에 대한 보다 정확하고 풍부한 출력을 제공

➤ 내부적 회귀 알고리즘을 통해 최적의 문제해결 결과를 도출할 때 까지 실행을 반복할 수 있음

➤ AGI의 핵심 알고리즘으로 예상

➤ **Function (x) Agent(O)**

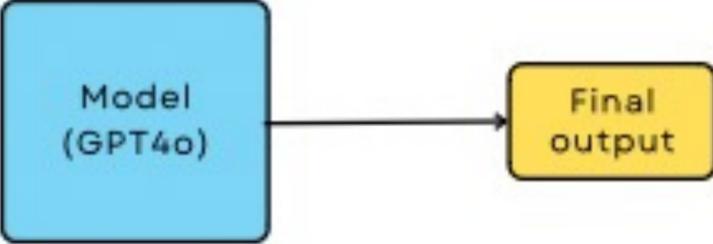
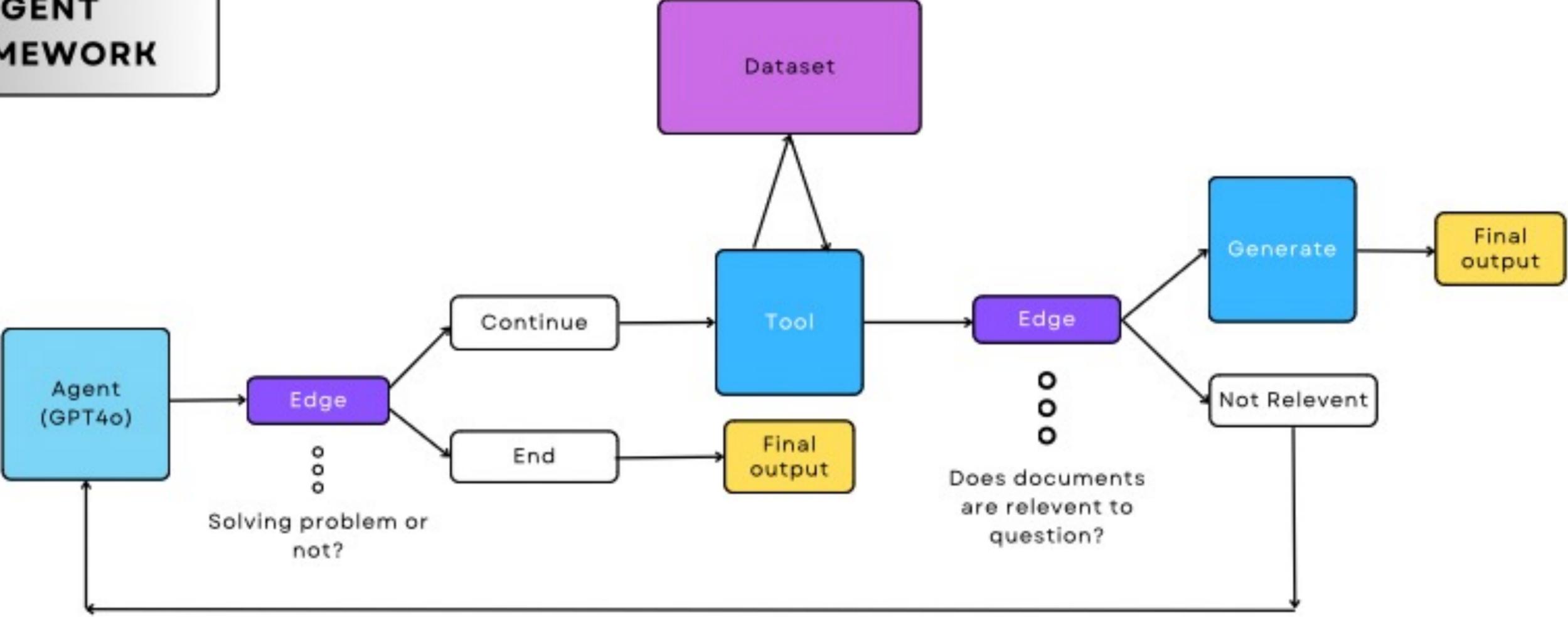
문제 정의:

미적분학 1 수업 당시의 문제 상황:

1. 학생들의 사고력 증진을 위해 문제 해설이 제공되지 않는 과제가 존재
2. 이에 대한 해설을, week test 시간에 TA들이 진행
3. 분반 별로 TA 수가 일정하기 때문에, 모든 학생들에게 적절한 솔루션을 제공하기 어려움

+가정 1. GPT4는 개념 설명에는 강력하지만, 연역적 추론 즉 문제풀이에는 약한 모습을 보인다.

AGENT FRAMEWORK



Components:

노드 (Nodes):

Agent:

사용자의 입력을 처리하는 노드
사용자가 입력한 질문이나 문제를 받아들여, 이를 분석한 뒤 'should solve problem' 엣지를 통해 판단을 진행

Tool:

벡터 데이터베이스와 상호작용하여 유사한 문제를 검색하고, 그 문제와 관련된 힌트나 풀이 과정을 탐색
학생이 풀어갈 수 있도록 적절한 힌트나 단서를 제공하는 방식으로 접근

Generate:

문제가 필요할 때마다 학생에게 적절한 힌트나 다음 단계의 풀이 과정을 제시하는 역할
정답을 바로 제공하지 않고, 풀이 과정에서 필요한 개념이나 방법을 이해할 수 있도록 도움을 줌

Rewrite:

만약 검색된 문제가 적절하지 않거나, 힌트가 충분하지 않다고 판단되면, 입력된 질문을 재구성하여 새로운 방향으로 유도하는 역할

Components:

엣지 (Edges):

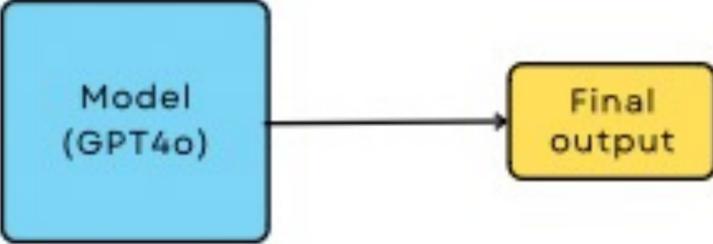
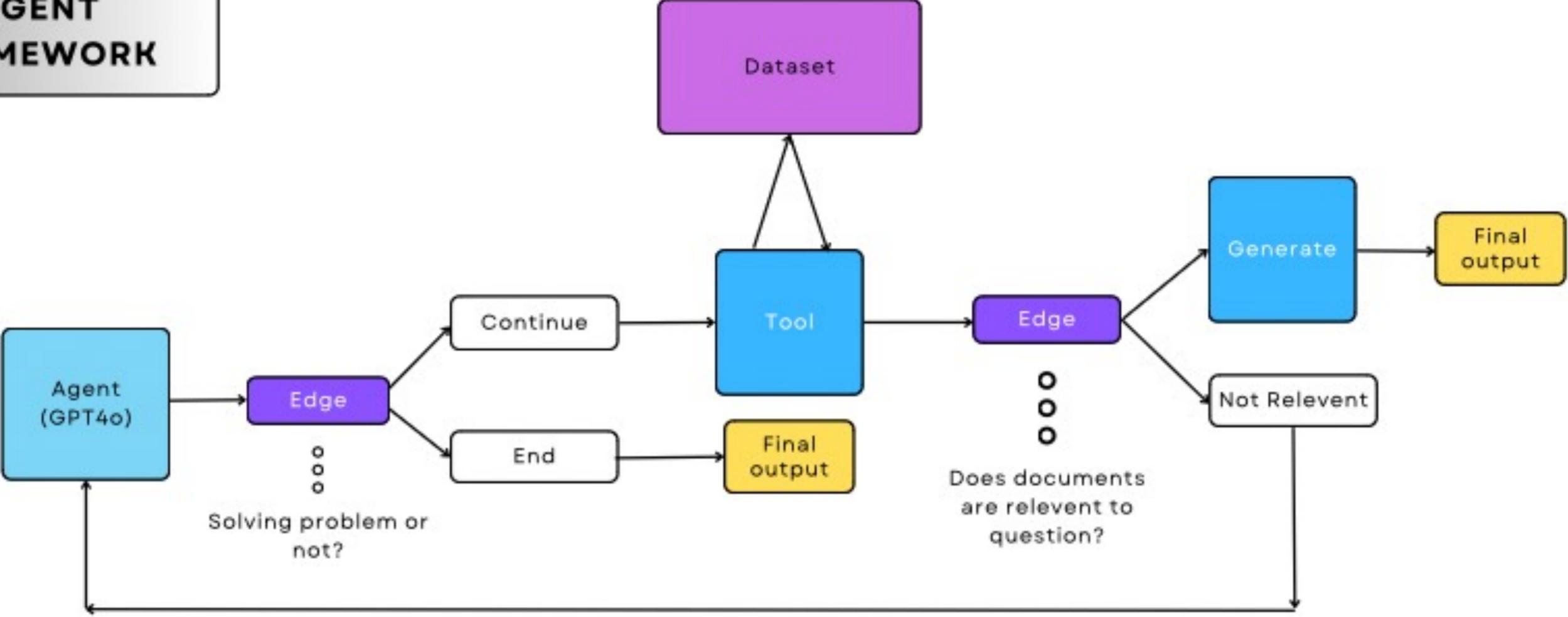
should solve problem:

이 엣지는 Agent가 사용자의 입력을 분석하여, 해당 입력이 문제 풀이를 요구하는지 아닌지를 판단하는 역할
문제 풀이가 필요하지 않다면, 개념 설명이나 간단한 힌트를 제공하고 종료. 문제 풀이가 필요하다면, Tool 노드로 전달하여 유사한 문제를 검색

relevant:

이 엣지는 Tool 노드에서 검색된 유사한 문제나 풀이가 현재 문제 해결에 적합한지를 판단.
적합하다고 판단되면, Generate 노드로 전달하여 학생이 스스로 문제를 풀어나갈 수 있도록 적절한 힌트와 단계를 제공
적합하지 않다면, Rewrite 노드를 통해 질문을 재구성하거나 새로운 접근 방식을 제안.

AGENT FRAMEWORK



Components:

3. 시스템 동작 흐름

Agent 입력 처리:

사용자가 수학 문제나 개념 설명을 요청하면 Agent가 이를 받아들여 'should solve problem' 엡지를 통해 분석
문제 풀이가 필요하지 않다면, 간단한 개념 설명이나 참고할 수 있는 자료를 제공.

문제 풀이 필요 판단:

문제 풀이가 필요하다고 판단되면, 문제를 Tool 노드로 전달하여 유사한 문제나 풀이 과정을 검색.

Tool 벡터 검색 및 힌트 제공:

Tool은 벡터 데이터베이스에서 유사한 문제를 검색하고, 그 문제의 풀이 과정에서 유용한 힌트나 단계를 추출.

'relevant' 엡지로 적합성을 판단한 후, Generate 노드로 전달.

적합성 검사:

'relevant' 엡지는 검색된 힌트나 풀이 과정이 현재 문제 해결에 도움이 되는지를 판단합니다.
적합한 힌트나 풀이 단계가 발견되면, Generate 노드에서 이를 바탕으로 학생이 스스로 풀어갈 수 있도록 단계별로 힌트를 제공.

Components:

3. 시스템 동작 흐름

•힌트 제공 및 유도:

- Generate 노드는 학생이 문제를 해결할 수 있도록 힌트를 단계별로 제공하며, 필요시 추가적인 설명을 덧붙임
- 정답을 직접 제시하는 것이 아니라, 풀이의 방향성을 제시하여 학생이 문제 해결 과정을 스스로 탐구하도록 유도

•질문 재구성 (필요시):

- 만약 제공된 힌트가 충분하지 않거나 적절하지 않다고 판단되면, Rewrite 노드를 통해 질문을 재구성하거나 새로운 접근 방식을 제안
- 새롭게 구성된 질문은 다시 Agent로 전달되어 과정을 반복

4. Question

4. 동작 예시

1. 사용자가 "미적분 문제 $\int(2x^3 - x^2 + 3) dx$ 를 풀어줘"라고 입력.
 2. Agent는 문제 풀이가 필요하다고 판단하고, 문제를 Tool 노드로 전달.
 3. Tool 노드는 유사한 문제를 벡터 데이터베이스에서 검색하고, 그 문제의 풀이 과정에서 유용한 힌트를 추출
 4. 'relevant' 엡지가 해당 힌트가 적절하다고 판단하면, Generate 노드에서 학생이 스스로 문제를 풀어갈 수 있도록 단계별 힌트를 제공
 5. 예시:
 - "우선 이 적분에서 각각의 항을 분리해보세요. 적분의 기본 법칙에 따라 각 항을 적분하는 방법을 생각해 보세요."
 - "적분 상수를 잊지 말고 고려하세요. 각 항을 적분한 후에는 상수를 추가해야 합니다."
- 필요시 추가 힌트를 제공

4. Question

4. 동작 예시

한 마디로 Agent 기능을 요약하자면:

인공지능에게 solution 상황에 가장 올바른 맞춤형 데이터 기반 프롬프트를 입력하자

>> 1. 데이터셋 구성 / 2.Prompt algorithm

즉 hard prompt와 관련된 문제임.

[] 코딩을 시작하거나 #로 코드를 생성하세요.

```

# Add comments here for code explanation
from crewai import Agent, Task, Crew

general_agent = Agent(role = "Math TA",
                      goal = ""Step by step, Provide the ideas or hints to the students that are asking mathematical questions and you should not give them the answer.your role is to help student to solve, not just solving problem"",
                      backstory = ""You are an excellent math TA that likes to solve math questions in a way that everyone can understand your solution"",
                      allow_delegation = False,
                      verbose = True,
                      llm = llm)

task = Task(description=""how to solve the integration? :
x
3
+3x
2
+3x+1
x
2
+2x+1
dx"",
            agent = general_agent,
            expected_output=""A hint or idea.not the answer"")

crew = Crew(
    agents=[general_agent],
    tasks=[task],
    verbose=True
)

result = crew.kickoff()

print(result)

```

```

WARNING:opentelemetry.trace:Overriding of current TracerProvider is not allowed
[DEBUG]: == Working Agent: Math TA
[INFO]: == Starting Task: how to solve the integration:

```



활동을 통해 얻고자 하는 것

FIRST IN CHANGE

Components:

1. 교내 자체 이용가능 모델 개발 파이프라인 구축
2. 교내 자체 이용가능 solution 파이프라인 구축
 - 공공기관에서 인공지능 도입 사례 선도, 공공기관에서의 인공지능 도입 촉진

FIRST IN CHANGE





THANK YOU

FIRST IN CHANGE