

Re-TEST ; 복습용 시험지 및 해설지 생성 시스템

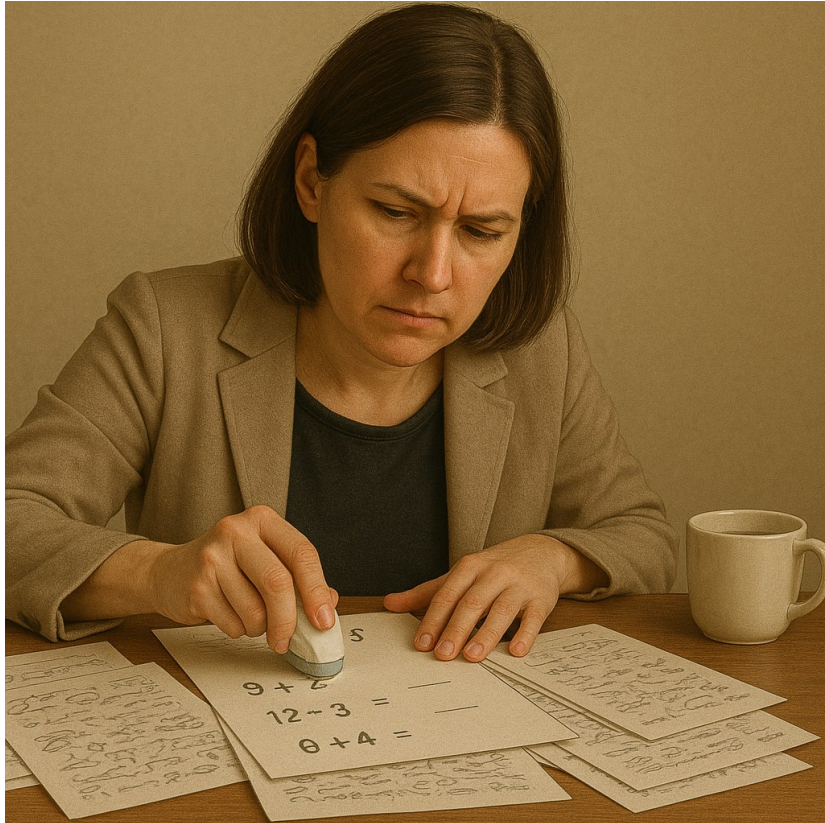
2025 Capstone Design II

문창집 2020270647

dwhouse0@korea.ac.kr

고려대(세종) 컴퓨터융합소프트웨어학과





Pictured by OpenAI

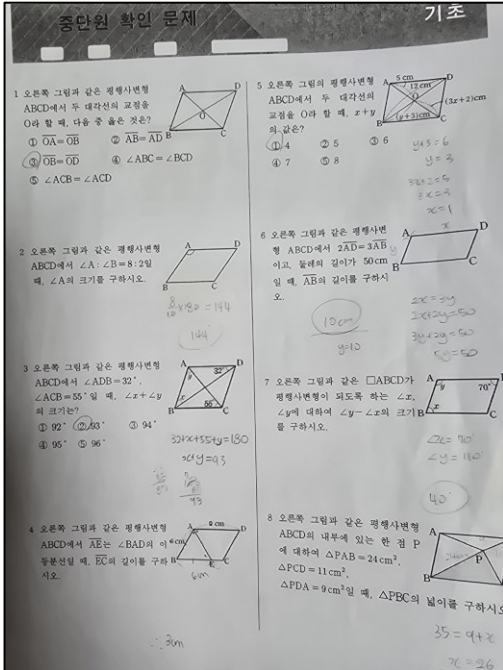
Digital ↔ Analog 학습의 혼합 시대에서,
여전히 종이 문제집 / 시험지를 많이 사용하고 있음.

복습 및 반복학습 시

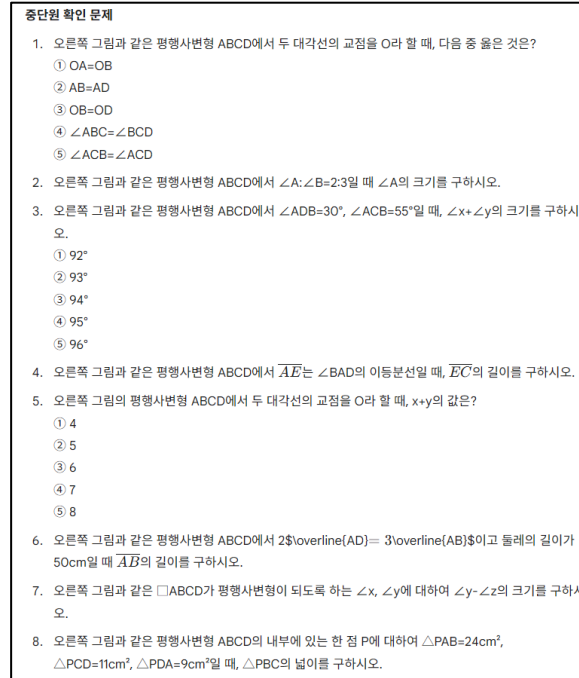
직접 쓴 문제지/시험지를 재활용하는 것은 어려우며,
문제를 다시 구성하고 정답과 해설을 작성하는 시간이 많이 소요됨.

 Target

- 반복학습의 주체인 학생
- 수업 자료를 재활용하는 교사 또는 강사



필기 / 낙서로 지저분해진 시험지



문제의 TEXT만 정확히 추출!

ReTest;

복습용 시험지 및 해설지 생성 시스템

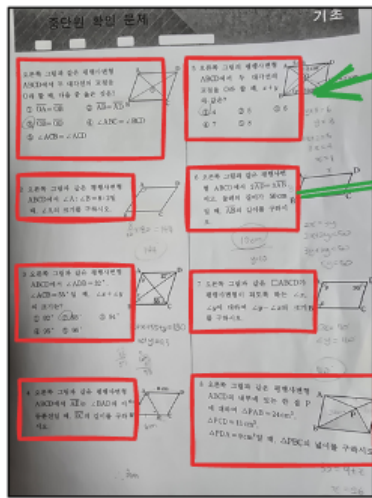
프로젝트 목표

필기된 시험지 이미지에서 문제를 추출하고,
새로운 복습용 시험지와 해설지를 자동 생성

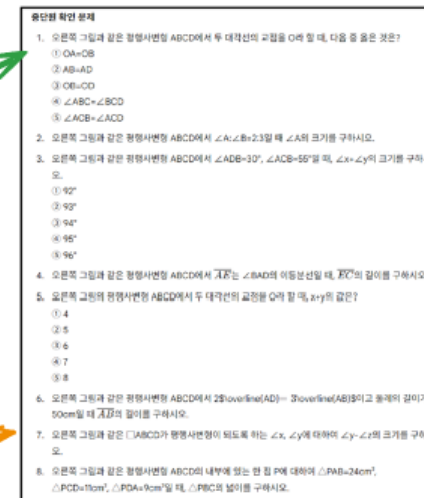
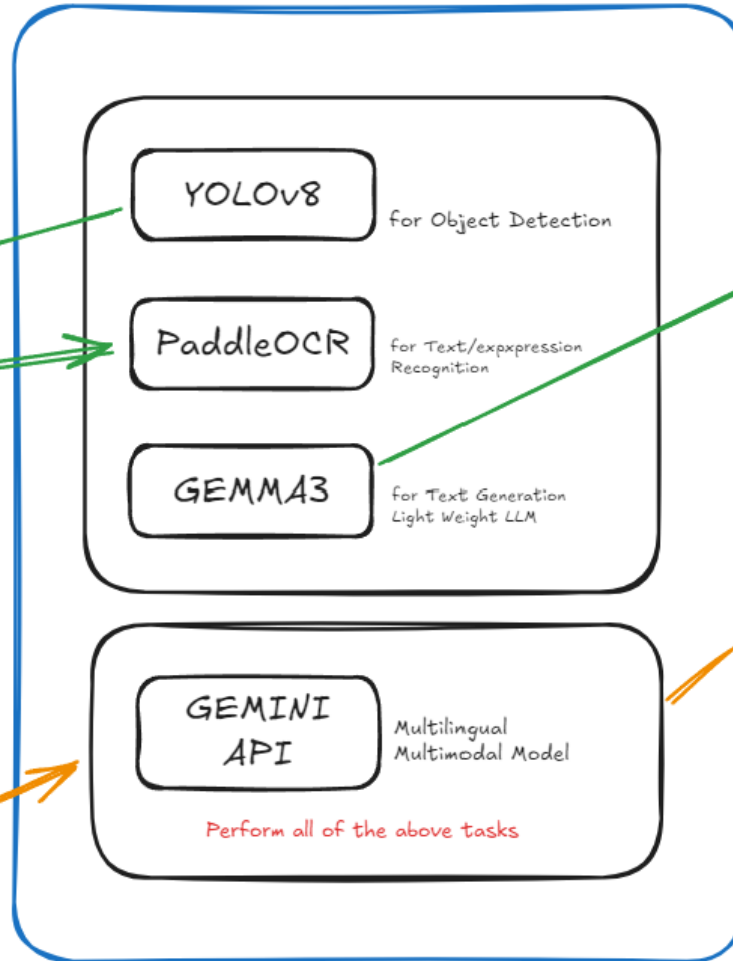
배경 및 필요성

- 디지털과 아날로그 학습의 혼합
- 반복학습 효율성 향상
- 교사의 해설 작성 부담 완화

RE-TEST



Dirty Paper



Clean Text

원래 계획 VS 실제 구현

시스템 구성

- 1st. 문제 영역 탐지(필기 제외) (with YOLO)
- 2nd. OCR, 수식 인식(with DONUT / Paddle..)
- 3rd. 문제 선택 (in FE) → 시험지 구성 (with LLM)
- 4th. Local LLM을 활용한 시험지/해설지 생성
- 5th. Streamlit을 활용한 웹 UI 구현

실제 구현

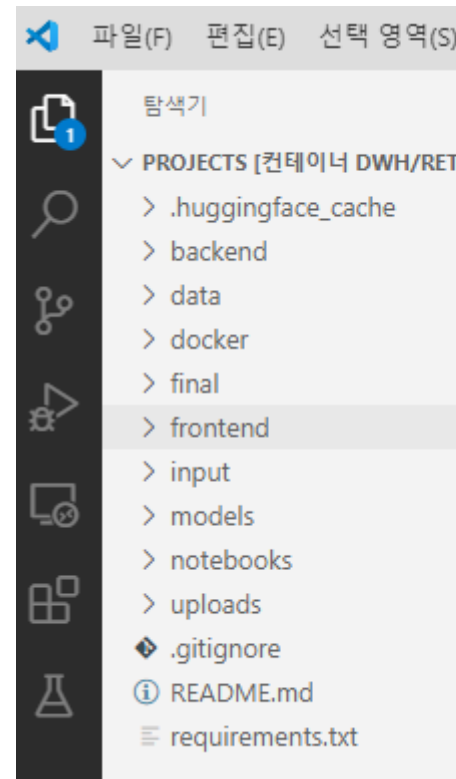
- 1st. Google Gemini API 연결
- 2nd. Prompt Engineering
- 3rd. Streamlit을 활용한 웹 UI 구현
- 4th. LLM을 활용한 시험지/해설지 생성

Dev Environment

1. Linux 20.04 + 3090(VRAM12GB)*4 Server
2. Docker Container Configure
3. Hugging Face / for Gemma3 Model
4. Ultralytics / YOLOv8 – Bbox Object Detection
5. Google AI Studio / for Gemini 2.5 Flash API
6. Google Cloud / for GPU inference
7. Streamlit / for Web UI Implementation



Repository Construction



```

dwh@icps2:~$ nvidia-smi
Fri May 23 03:06:08 2025
-----
NVIDIA-SMI 550.78      Driver Version: 550.78      CUDA Version: 12.4
-----+-----
GPU  Name        Persistence-M  Bus-Id  Disp.A  Volatile Uncorr. ECC
 Fan  Temp    Perf      Pwr:Usage/Cap  Bus-Id  Memory-Usage  GPU-Util  Compute M.
                                Pwr:Usage/Cap  Memory-Usage  GPU-Util  MIG M.
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
  0  NVIDIA GeForce RTX 3090      Off      00000000:19:00.0 Off      0%      N/A
    0%   35C    P8          35W / 370W      13MiB / 24576MiB      0%      Default
                                           N/A
  1  NVIDIA GeForce RTX 3090      Off      00000000:1A:00.0 Off      0%      N/A
    0%   36C    P8          37W / 370W      13MiB / 24576MiB      0%      Default
                                           N/A
  2  NVIDIA GeForce RTX 3090      Off      00000000:67:00.0 Off      0%      N/A
    0%   40C    P8          32W / 370W      416MiB / 24576MiB      0%      Default
                                           N/A
  3  NVIDIA GeForce RTX 3090      Off      00000000:68:00.0 Off      0%      N/A
    0%   39C    P8          30W / 370W      80MiB / 24576MiB      0%      Default
                                           N/A
-----+-----
Processes:
GPU  GI  CI  PID  Type  Process name  GPU Memory Usage
 ID  ID  ID
-----+-----+-----+-----+-----+-----+-----
  0  N/A  N/A  1319  G    /usr/lib/xorg/Xorg  4MiB
  1  N/A  N/A  1319  G    /usr/lib/xorg/Xorg  4MiB
  2  N/A  N/A  1319  G    /usr/lib/xorg/Xorg  4MiB
  2  N/A  N/A  2288939  C    ../.pyenv/versions/3.9.19/bin/python3  400MiB
  3  N/A  N/A  1319  G    /usr/lib/xorg/Xorg  56MiB
  3  N/A  N/A  1510  G    /usr/bin/gnome-shell  11MiB
-----+-----
dwh@icps2:~$ docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED          STATUS          PORTS
83c21383ff25   dwh/retest    "/opt/nvidia/nvidia_..." 2 days ago      Up 2 days      0.0.0.0:894->8894/tcp, :::8894->8894/tcp, 0.0.0.0:32768->8501/tcp, :::32768->8501/tcp
dwh_container2

```

Server Environment & Docker Process

Datasets (1) – For YOLO, OCR Model Tuning

$x > 0$

이므로 아래와 같다. 이때 y

$\int_0^1 g(x) dx$ 1298

$\sin t(2\cos t - 1) = 0$ 적분

좌표를 1305

```

1  {
2  "id": "901475c4-4fe7-48dd-ac59-4d63a90043c8",
3  "chapter_id": "1217",
4  "material_type": "인쇄체",
5  "segments": [
6    {
7      "box": [
8        [6, 8],
9        [54, 8],
10       [54, 27],
11       [6, 27]
12      ],
13      "type": "수식/텍스트",
14      "type_detail": "수식/텍스트",
15      "equation": "$\displaystyle x > 0$"
16    },
17    {
18      "box": [
19        [6, 48],
20        [156, 48],
21        [156, 71],
22        [6, 71]
23      ],
24      "type": "수식/텍스트",
25      "type_detail": "수식/텍스트",
26      "equation": "이므로 아래와 같다. 이때 $\displaystyle y$"
27    },
28    {
29      "box": [
30        [6, 91],
31        [135, 91],
32        [135, 129],
33        [6, 129]
34      ],
35      "type": "수식/텍스트",
36      "type_detail": "수식",
37      "equation": "$\displaystyle \int_0^1 g(x) dx$ $\displaystyle 1298$"
38    },
39  ],
40  "type": "수식/텍스트",
41  "type_detail": "수식/텍스트",
42  "equation": "$\displaystyle \sin\{t\}(2\cos\{t\}-1) = 0$ 적분"
43  },
44  {
45    "box": [
46      [6, 149],
47      [206, 149],
48      [206, 172],
49      [6, 172]
50    ],
51    "type": "수식/텍스트",
52    "type_detail": "수식/텍스트",
53    "equation": "$\displaystyle \sin\{t\}(2\cos\{t\}-1) = 0$ 적분"
54  },
55  {
56    "box": [
57      [6, 192],
58      [107, 192],
59      [107, 213],
60      [6, 213]
61    ],
62    "type": "수식/텍스트",
63    "type_detail": "수식/텍스트",
64    "equation": "좌표를 $\displaystyle 1305$"
65  }
66  ]
67  }

```

	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
메타데이터 정보 (다중기입가능)	영상이미지	이미지	인쇄체 38,997건 손글씨 149,268건	png	json	세그멘테이션이미지, LaTeX라벨링(수식/텍스트), 구성정보 라벨링(도형)
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체 수집	2021년	(주)씨유박스	(주)클라우드웍스, (주)제제듀	(주)클라우드웍스, (주)제제듀	
	데이터 문의처	기관명	문의담당명	전화번호 (유선전화번호기입)	메일주소	
		(주)씨유박스	김신	02-6277-7872	skim@cubox.aero	
	데이터 소개	인쇄체, 손글씨 수식 등의 인식 정밀도를 향상시키기 위해 수학 교과과정에서 사용 되는 수식, 도형, 풀이 및 첨삭 낙서기호 등에 대한 OCR 데이터 구축				
주요키워드	OCR, Mathematical Expression Recognition, LaTeX generation, Handwritten Mathematical Expression recognition, 수식 인식, 도형 인식, 손글씨 인식					

Datasets (2), (3) – For YOLO, OCR Model Tuning

두 다항식 $A = 3x^2 - 4xy + y^2$, $B = 2x^2 - xy + 5y^2$ 에 대하여 $2A - B$ 를 계산하여 x, y 에 대한 식으로 나타내시오.

$$\begin{aligned} 2A - B &= 2(3x^2 - 4xy + y^2) - (2x^2 - xy + 5y^2) \\ &= 6x^2 - 8xy + 2y^2 - 2x^2 + xy - 5y^2 \\ &= 4x^2 - 7xy - 3y^2 \end{aligned}$$

Handwritten solution for $2A - B$ showing the algebraic steps:

$$\begin{array}{r} 2A - B \\ 6x^2 - 8xy + 2y^2 \\ + -2x^2 + xy - 5y^2 \\ \hline 4x^2 - 7xy - 3y^2 \end{array}$$

```
"question_info": [
  {
    "question_grade": "H",
    "question_term": 1,
    "question_unit": "01",
    "question_topic": "0101001",
    "question_topic_name": "다항식의 덧셈과 뺄셈",
    "question_type1": "서술",
    "question_type2": 2,
    "question_condition": 0,
    "question_sector1": "계산",
    "question_sector2": "변화와 관계",
    "question_step": "기본",
    "question_difficulty": 3,
    "question_contents": 0,
    "question_rtime": 73,
    "question_success_rate": 94.53
  }
],
"OCR_info": [
  {
    "figure_text": null,
    "question_text": "두 다항식 $A=3x^2-4xy+y^2$, $B=2x^2-xy+5y^2$에 대하여 $2A-B$를 계산하여 $x$, $y$에 대한 식으로 나타내시오.",
    "question_bbox": [
      {
        "answer_filename": "H_1_01_25766_84186_A.png",
        "answer_text": "$2A-B$ $$$$(3x^2-4xy+y^2)-(2x^2-xy+5y^2)$ $$$6x^2-8xy+2y^2-2x^2",
        "answer_bbox": [
          {
            "type": "line",
            "bbox": [
              1,
              1,
              55,
              18
            ]
          }
        ],
        "text": null
      }
    ]
  }
]
```

Screenshot of a Jupyter Notebook interface showing a file explorer and a data table:

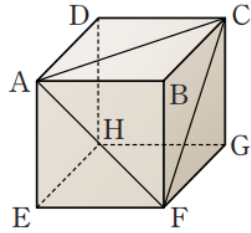
File Explorer: `H_1_01_25763_84170.txt` × `02_solution1_tuning.ipynb`

Navigation: `data > 03_generation_data > train > labels > Question > hs_common > H_1_01_25763_84170.txt`

1	0.000000	0.443144	0.223077	0.882943	0.415385
---	----------	----------	----------	----------	----------

Datasets (4) – For YOLO, OCR Model Tuning

016 오른쪽 그림과 같은 직육면체 ABCD-EFGH의 모든 모서리의 길이의 합이 24이고, 삼각형 AFC의 세 변의 길이의 제곱의 합이 28일 때, 이 직육면체의 겉넓이를 구하시오.



$\overline{AB}=a$, $\overline{BC}=b$, $\overline{BF}=c$ 로 놓으면 직육면체의 모든 모서리의 길이의 합은

$$4(a+b+c)=24$$

$$\therefore a+b+c=6 \quad \text{..... ㉠}$$

삼각형 AFC의 세 변의 길이의 제곱의 합은 피타고라스 정리를 이용하면

$$(a^2+b^2)+(b^2+c^2)+(c^2+a^2)=28$$

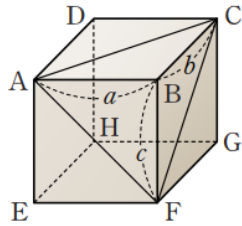
$$\therefore a^2+b^2+c^2=14 \quad \text{..... ㉡}$$

이때 직육면체의 겉넓이는 $2(ab+bc+ca)$ 이므로

$$2(ab+bc+ca)=(a+b+c)^2-(a^2+b^2+c^2)$$

$$=6^2-14=22$$

22



```
"learning_data_info": [
  {
    "class_num": 1,
    "class_name": "문항(텍스트)",
    "class_info_list": [
      {
        "Type": "Bounding_Box",
        "Type_value": [
          [
            [
              64,
              18,
              36,
              23],
            [
              "class_num": 2,
              "class_name": "해설(텍스트)",
              "class_info_list": [
                {
                  "Type": "Bounding_Box",
                  "Type_value": [
                    [
                      17.83,
                      294.96,
                      282.91,
                      484.52
                    ]
                  ],
                  "text_description": "$\\overline{\\mathrm{AB}}=a, \\overline{\\mathrm{BC}}=b, \\overline{\\mathrm{BF}}=c$로 놓으면 직"
                }
              ]
            },
            {
              "Type": "Bounding_Box",
              "Type_value": [
                [
                  17.41,
                  493.18,
                  428.42,
                  690.6
                ]
              ],
              "text_description": "제곱의 합은 피타고라스 정리를 이용하면 \\[ \\begin{array}{l} \\left(a^2+b^2\\right)+\\left(b^2+c^2\\right)+\\left(c^2+a^2\\right)=28 \\end{array} ]"
            }
          ]
        ]
      }
    ]
  }
]
```



Data Preparation

YOLO Class Mapping:

- 문항(텍스트) -> question (class_id: 0)
- 문항(이미지) -> question (class_id: 0)
- 정답(텍스트) -> answer (class_id: 1)
- 정답(이미지) -> answer (class_id: 1)
- 오답(텍스트) -> Excluded
- 오답(이미지) -> Excluded
- 해설(텍스트) -> Excluded
- 해설(이미지) -> Excluded

JSON path: /projects/data/04_Problem/Test/labels/S3_중등_2_004108.json
 Image path: /projects/data/04_Problem/Test/images/S3_중등_2_004108.png
 Image exists: True
 Converted label count: 2
 First label: [0, 0.4165912518853695, 0.11630434782608697, 0.6784313725490196, 0.074637681

3 다음 중 일차함수를 가장 바르게 나타낸 것은?

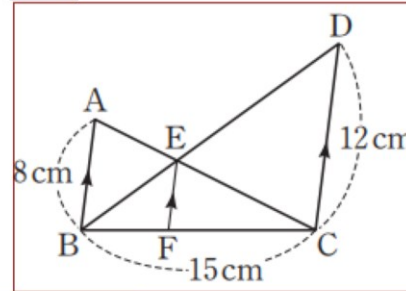
- ① $y = ax$
- ② $y = ax + b$
- ③ $y = ax + b$ (a, b 는 상수)
- ④ $y = ax + b$ (a, b 는 상수, $a \neq 0$)
- ⑤ $y = ax + b$ (a, b 는 상수, $b \neq 0$)

3 일차식이 되려면 x 항의 계수가 0이 아니어야 한다.

따라서 일차함수 $y = ax + b$ (a, b 는 상수, $a \neq 0$)

3 ④

03 오른쪽 그림에서 $\overline{AB} \parallel \overline{EF} \parallel \overline{DC}$ 일 때, \overline{FC} 의 길이를 구하시오.



$\triangle ABE \sim \triangle CDE$ (AA 대응) 이므로
 $\overline{BE} : \overline{DE} = \overline{AB} : \overline{CD} = 8 : 12 = 2 : 3$ 이다.
 또 $\triangle BEF \sim \triangle BDC$ (AA 대응) 이므로
 $\overline{BC} : \overline{FC} = \overline{BD} : \overline{ED} = (2+3) : 2$
 $5 : \overline{FC} = 5 : 3$, $5\overline{FC} = 45$ 이다.
 따라서 $\overline{FC} = 9$ cm이다.

어느 제과점에서 하루에 12.4l kg의 밀가루 반죽을 0.35 kg씩 그릇에 나누어 담아 식빵을 만들려고 합니다. 버리는 양이 없이 모두 그릇에 반죽을 담는다면 그릇이 몇 개 필요하고, 또 가장 적게 들어갈 그릇에는 몇 kg이 들어가는지 구해 보세요.

$$\begin{array}{r} 35 \\ 0.35 \overline{) 12.41} \\ \underline{105} \\ 191 \\ \underline{175} \\ 0.16 \end{array}$$



예) $12.41 \div 0.35 = 35.45 \dots$ 입니다. 나머지 양도 담아야 하므로 그릇은 36개 필요합니다.

35개의 그릇에는 0.35 kg씩 담고 남은 밀가루 반죽은 나머지 그릇에 담으면 됩니다.

(35개의 그릇에 들어간 밀가루 양) $= 0.35 \times 35 = 12.25$ (kg)

(가장 적게 들어간 그릇의 밀가루 양) $= 12.41 - 12.25 = 0.16$ (kg)

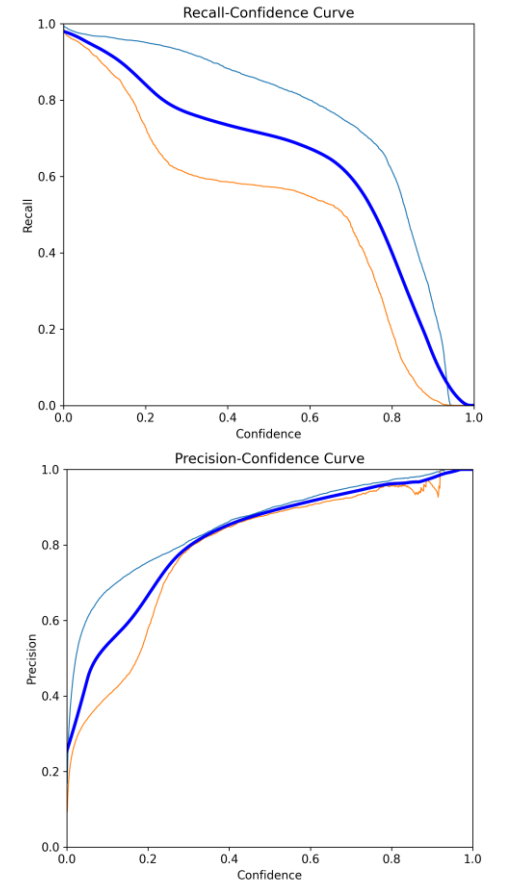
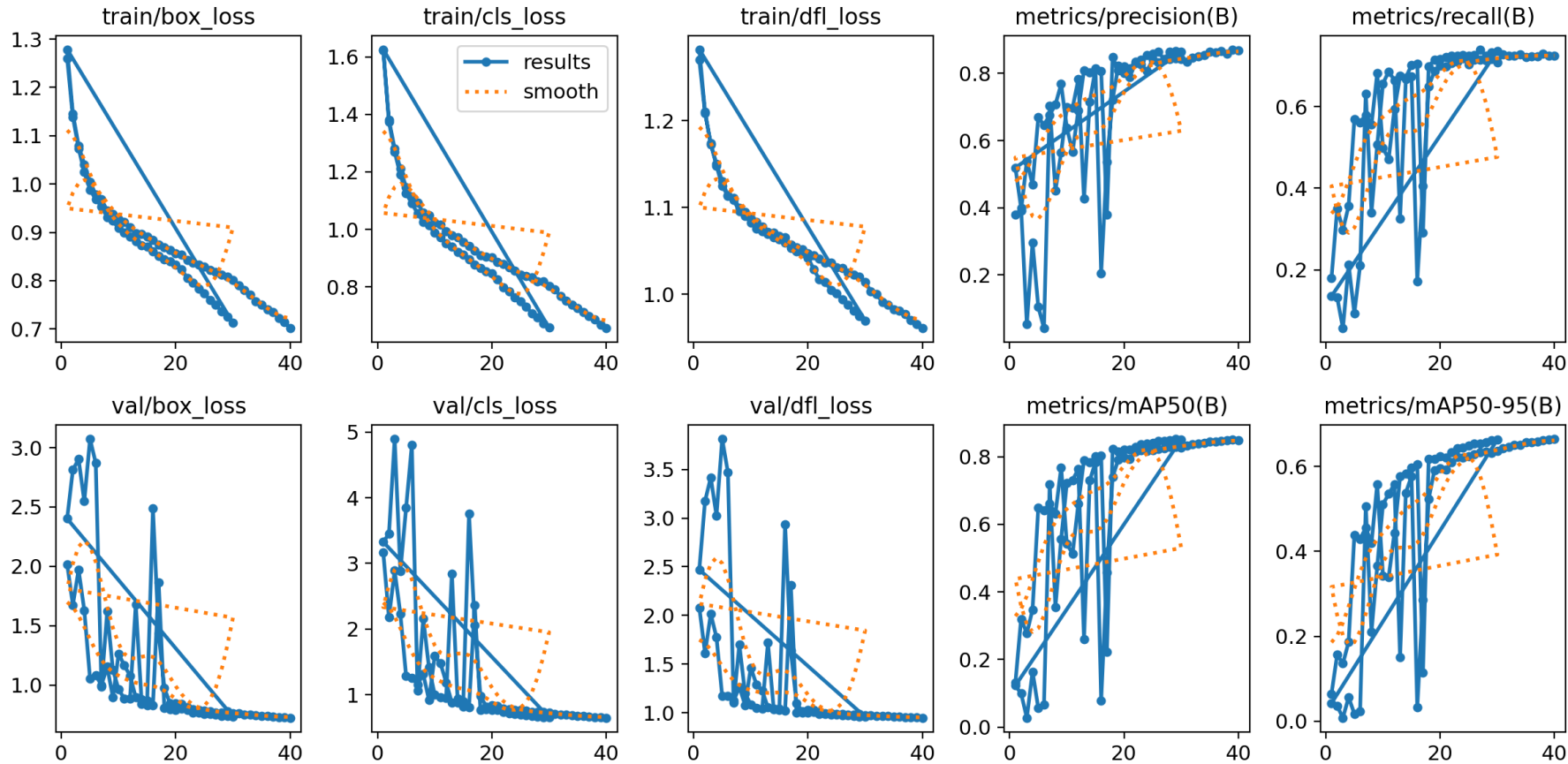
답 36개, 0.16 kg

Train 데이터 변환 중...
 Processing train: 100% | 12796/12796 [00:04:00:00, 2826.29it/s]
 - 처리됨: 12795
 - 건너뛴: 0
 - 오류: 1

Valid 데이터 변환 중...
 Processing val: 100% | 2742/2742 [00:01:00:00, 2704.22it/s]
 - 처리됨: 2742
 - 건너뛴: 0
 - 오류: 0

Test 데이터 변환 중...
 Processing test: 100% | 2742/2742 [00:00:00:00, 2910.37it/s]
 - 처리됨: 2742
 - 건너뛴: 0
 - 오류: 0

YOLO Training Result




YOLO Training Result

S3_00_1_006038.png...

2 question 0.86 해하시오.
 question 0.57
 (1) 63
 $(1) 63 = 3^2 \times 7$
 answer 0.83
 $3^2 \times 7$

S3_00_2_019277.png...

question 0.86



question 0.86

① 36°, 72°, 72°
 ② 60°, 60°, 60°
 ③ 40°, 50°, 50°
 ④ 108°, 36°, 36°
 ⑤ 80°, 50°, 50°

answer 0.86

예) $\angle B = \angle C$ 이고 $\angle B = 2\angle A$ 이므로 $\angle C = 2\angle A$
 즉 $\angle A + \angle B + \angle C = 5\angle A = 180^\circ$ 이므로
 $\angle A = 36^\circ$, $\angle B = \angle C = 72^\circ$

S3_00_3_012473.png...

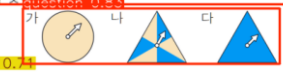
3~4
 4
 question 0.83
 question 0.83
 8L 600 mL
 - 7L 450 mL
 answer 0.69
 1 L 150 mL

S3_00_5_010094.png...

question 0.89

question 0.89를 하고 있습니다. 화살이 주황색에 멈출 가능성이 가장 낮은 회전판을 찾아주세요.

question 0.89



answer 0.74

예)가: 전체가 하늘색 \Rightarrow 불가능하다 $\frac{0}{6}$
 나: 전체 6칸 중의 3칸 \Rightarrow $\frac{3}{6} = \frac{1}{2}$
 다: 전체가 주황색 \Rightarrow 확실히 $\frac{1}{3}$
 따라서 $0 < \frac{1}{3} < \frac{1}{2}$ 이므로 화살이 주황색에 멈출 가능성이 가장 낮은 회전판은 가입니다.

S3_00_1_007255.png...

question 0.87

7. 임의의 양의 실수 a, b, c, d 에 대하여 $ax^3 + bx^2 + cx + d = 0$ 의 한 근을 α ($\alpha \neq 0$)라고 할 때, 다음 중 옳은 것을 고르시오.

가. $-\alpha$ 는 $ax^3 - bx^2 + cx - d = 0$ 의 근이다.
 나. $\frac{1}{\alpha}$ 은 $dx^3 + cx^2 + bx + a = 0$ 의 근이다.
 다. $-\frac{1}{\alpha}$ 은 $dx^3 - cx^2 + bx - a = 0$ 의 근이다.

7. 답 가, 나, 다

가. $ax^3 - bx^2 + cx - d = 0$ 에 $x = -\alpha$ 를 대입
 하면 $-a\alpha^3 - b\alpha^2 - c\alpha - d = 0$
 $a\alpha^3 + b\alpha^2 + c\alpha + d = 0$ 이므로 참

나. $dx^3 + cx^2 + bx + a = 0$ 에 $x = \frac{1}{\alpha}$ 을 대입
 하면 $\frac{d}{\alpha^3} + \frac{c}{\alpha^2} + \frac{b}{\alpha} + a = 0$
 $a\alpha^3 + b\alpha^2 + c\alpha + d = 0$ 이므로 참

다. $dx^3 - cx^2 + bx - a = 0$ 에 $x = -\frac{1}{\alpha}$ 를 대입
 하면 $-\frac{d}{\alpha^3} - \frac{c}{\alpha^2} + \frac{b}{\alpha} - a = 0$
 $a\alpha^3 + b\alpha^2 + c\alpha + d = 0$ 이므로 참

OCR @ Cropped Images

```
1 import easyocr
2 import numpy as np
3 from PIL import Image
4 import torch
5 import matplotlib.pyplot as plt
6
7 # Initialize EasyOCR with Korean and English support
8 reader = easyocr.Reader(['ko', 'en'], gpu=True)
9
10 print("✅ EasyOCR initialized successfully!")
11 print(f" - GPU: {torch.cuda.is_available()}")
12 print(f" - Languages: Korean, English")
```

✓ 1.9s

- ✓ EasyOCR initialized successfully!
- GPU: True
- Languages: Korean, English

image 1/1 /projects/data/04_Problem_YOLO/images/test/S3_중등_1_000021.png: 224x640 1 question, 29.6ms
Speed: 1.3ms preprocess, 29.6ms inference, 1.6ms postprocess per image at shape (1, 3, 224, 640)

YOLO Detections: S3_중등_1_000021.png

03 question 0.74
36과 60의 공약수의 개수는?

① 3 ② 4 ③ 5
④ 6 ⑤ 7

36과 60의 공약수는 최대공약수인 12의 약수이다. 따라서 36과 60의 공약수는 1, 2, 3, 4, 6, 12의 6개이다.

④

$$\begin{aligned} 36 &= 2^2 \times 3^2 \\ 60 &= 2^2 \times 3 \times 5 \\ \hline \text{(최대공약수)} &= 2^2 \times 3 = 12 \end{aligned}$$

Processing: S3_중등_1_000021.png
Found 1 detections

Question 1 (YOLO Conf: 0.745)

36과 60의 공약수의 개수는?

OCR on Question 1:
- 36과 60의 공약수의 개수는? (conf: 0.828)
- Combined: 36과 60의 공약수의 개수는?...
- Avg confidence: 0.828

Pipeline Test

YOLO Detections: test006.png

problem_0.85

25. 정규분포 $N(m, 2^2)$ 을 따르는 모집단에서 크기가 256인 표본을 임의추출하여 얻은 표본평균을 이용하여 구한 m 에 대한 신뢰도 95%의 신뢰구간이 $a \leq m \leq b$ 이다. $b-a$ 의 값은? (단, Z 가 표준정규분포를 따르는 확률변수일 때, $P(|Z| \leq 1.96) = 0.95$ 로 계산한다.) [3점]

- ① 0.49 ② 0.52 ③ 0.55 ④ 0.58 ⑤ 0.61

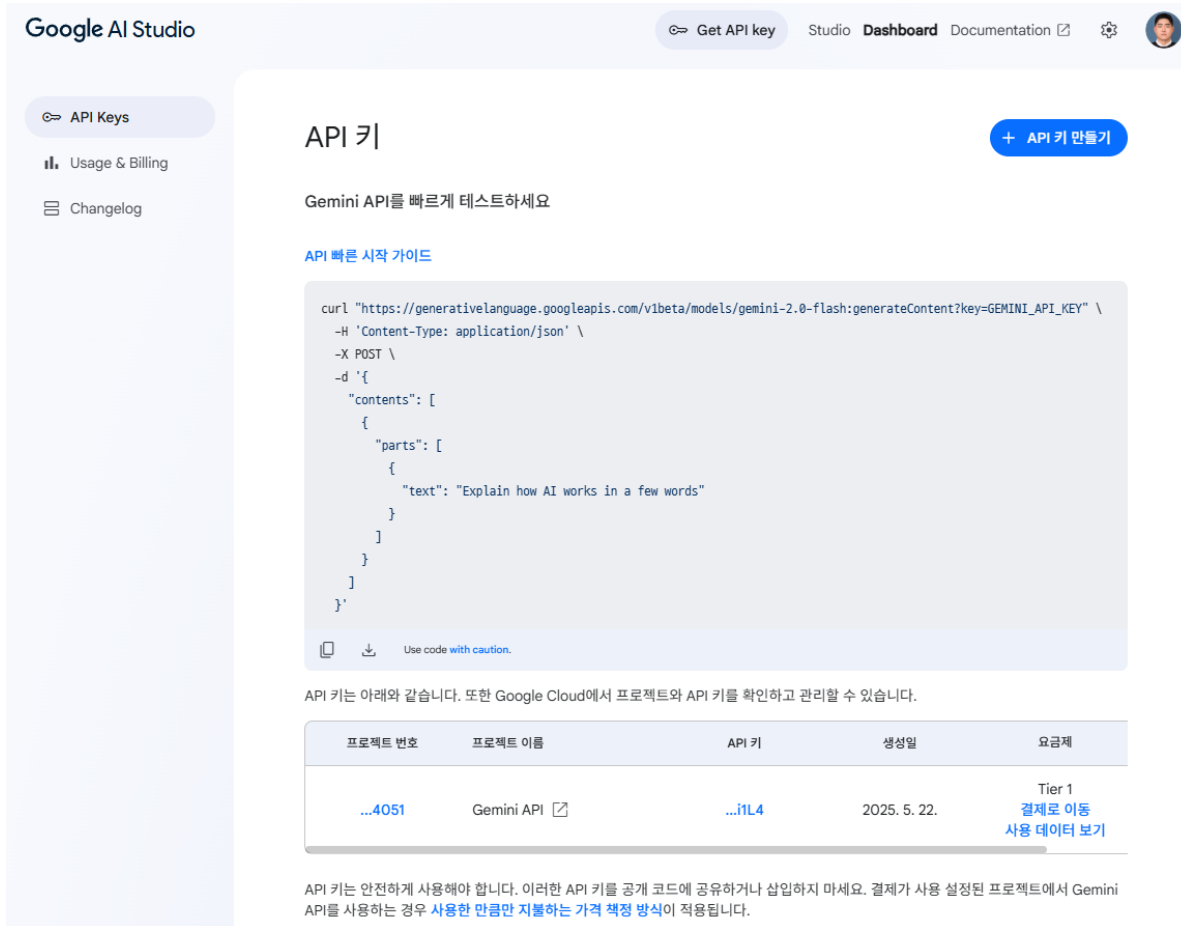
problem_0.93

26. 어느 학급의 학생 16명을 대상으로 과목 A와 과목 B에 대한 선호도를 조사하였다. 이 조사에 참여한 학생은 과목 A와 과목 B 중 하나를 선택하였고, 과목 A를 선택한 학생은 9명, 과목 B를 선택한 학생은 7명이다. 이 조사에 참여한 학생 16명 중에서 임의로 3명을 선택할 때, 선택한 3명의 학생 중에서 적어도 한 명이 과목 B를 선택한 학생일 확률은? [3점]

◆ Question 1: ----- 26. 어느 학급의 학생 16명을 대상으로 과목 A와 과목 B에 대한 선호도를 조사하였다. 이 조사에 참여한 학생은 과목 A와 과목 B 중 하나를 선택하였고, 과목 A를 선택한 학생은 9명, 과목 B를 선택한 학생은 - 명이다: 이 조사에 참여한 학생 16명 중에서 임의로 3명을 선택할 때, 선택한 3명의 학생 중에서 적어도 한 명이 과목 B를 선택한 학생일 확률은? [3점] [Confidence: 0.416]

◆ Question 2: ----- 25. 정규분포 $N(m, 2^2)$ 을 따르는 모집단에서 크기가 256인 표본을 임의 추출하여 얻은 표본평균을 이용하여 구한 m 에 대한 신뢰도 95%의 신뢰구간이 $a \leq m \leq b$ 이다: $b-a$ 의 값은? (단 Z 가 표준정규분포를 따르는 확률변수일 때, $P(|Z| \leq 1.96) = 0.95$ 로 계산한다) [3점] [Confidence: 0.472]

Gemini API



The screenshot shows the Google AI Studio interface. On the left, there are navigation tabs for 'API Keys', 'Usage & Billing', and 'Changelog'. The main content area is titled 'API 키' and includes a '+ API 키 만들기' button. Below the title, there is a text prompt: 'Gemini API를 빠르게 테스트하세요' and a link to 'API 빠른 시작 가이드'. A code block contains a cURL command for testing the Gemini API. Below the code, there is a table listing API keys with columns for '프로젝트 번호', '프로젝트 이름', 'API 키', '생성일', and '요금제'. The table shows one key for 'Gemini API' with ID '...i1L4', created on '2025. 5. 22.', and a 'Tier 1' pricing plan. A note below the table states that API keys should be used securely and that usage-based pricing applies.

```
curl "https://generativelanguage.googleapis.com/v1beta/models/gemini-1.5-flash:generateContent?key=GEMINI_API_KEY" \
-H 'Content-Type: application/json' \
-X POST \
-d '{
  "contents": [
    {
      "parts": [
        {
          "text": "Explain how AI works in a few words"
        }
      ]
    }
  ]
}'
```

프로젝트 번호	프로젝트 이름	API 키	생성일	요금제
...4051	Gemini API	...i1L4	2025. 5. 22.	Tier 1 결제로 이동 사용 데이터 보기



The screenshot shows the Google AI SDK installation and usage guide. It includes a table of contents on the left with sections like '시작하기', '모델 기능', '가이드', and '리소스'. The main content area is titled 'Google GenAI SDK 설치' and includes a code block for installing the SDK using pip. Below the code block, there is a section for '첫 번째 요청하기' with a code block for using the generateContent method. The code block shows how to import the genai module, create a client, and call the generate_content method with the model name and content.

```
pip install -q -U google-gemai
```

```
from google import genai

client = genai.Client(api_key="YOUR_API_KEY")

response = client.models.generate_content(
    model="gemini-1.5-flash", contents="Explain how AI works in a few words"
)

print(response.text)
```

Streamlit Application

Re-TEST;

Gemini Flash 2.0 API 기반
이미지로부터 수학 문제를 자동 인식하고,
복습용 시험지와 해설지를 생성하는 스마트 학습 도우미입니다.

TEST

4 + =
3 - x
5 +
8 ÷ =

AI

시작하기

INTRO PAGE

Re-TEST;

문제지 이미지 업로드

이미지에서 문제를 인식할 준비가 되었어요!

JPG, PNG 형식의 스캔된 시험지 혹은 복습하고 싶은 문항이 있는 사진을 업로드 해주세요.

시험지 이미지를 선택하여 업로드하거나 Drag&Drop 해주세요

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

위에 문제지 이미지를 올려주세요.

처음으로 돌아가기

시작하기 버튼 클릭 시

Streamlit Application

Re-TEST;

문제지 이미지 업로드

이미지에서 문제를 인식할 준비가 되었어요!

JPG, PNG 형식의 스캔된 시험지 혹은 복습하고 싶은 문항이 있는 사진을 업로드해주세요.

시험지 이미지를 선택하여 업로드하거나 Drag&Drop 해주세요

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

KakaoTalk_20250517_214633417.jpg 109.6KB

이미지 업로드 완료!

22 $\frac{1}{250}$ 입니다. ㉠에 알맞은 수 중에서 가장 작은 수는 얼마입니까?
 $\frac{1}{25 \times 10} = \frac{1}{5^3 \times 2 \times 2 \times 5} = \frac{2 \times 5}{5^4 \times 2^2} = \frac{2 \times 5}{(5^2 \times 2)^2}$
㉠ = 10.

업로드된 문제지 미리보기

문제 영역 인식하러 가기

처음으로 돌아가기

Image Uploaded

Re-TEST;

문제 영역 인식 중...

Gemini Flash 모델이 이미지 내 문제 문장을 추출하고 있습니다.

Gemini Flash가 문제지를 분석 중입니다...

시험지 이미지를 선택하여 업로드하거나 Drag&Drop 해주세요

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

KakaoTalk_20250517_214633417.jpg 109.6KB

이미지 업로드 완료!

22 $\frac{1}{250}$ 입니다. ㉠에 알맞은 수 중에서 가장 작은 수는 얼마입니까?
 $\frac{1}{25 \times 10} = \frac{1}{5^3 \times 2 \times 2 \times 5} = \frac{2 \times 5}{5^4 \times 2^2} = \frac{2 \times 5}{(5^2 \times 2)^2}$
㉠ = 10.

업로드된 문제지 미리보기

문제 영역 인식하러 가기

처음으로 돌아가기

Problem Recognition

Re-TEST;

문제 영역 인식 중...

Gemini Flash 모델이 이미지 내 문제 문장을 추출하고 있습니다.

Gemini 응답 원문 보기

응답 내용

```
```json
{
 "id": "problem_1",
 "text": "㉠/㉡x㉢ = 1/250 입니다. ㉠에 알맞은 수 중에서 가장 작은 수는 얼마입니까?"
}
```

문제 1개 인식 완료! 다음 단계로 진행하세요.

22  $\frac{1}{250}$  입니다. ㉠에 알맞은 수 중에서 가장 작은 수는 얼마입니까?  
 $\frac{1}{25 \times 10} = \frac{1}{5^3 \times 2 \times 2 \times 5} = \frac{2 \times 5}{5^4 \times 2^2} = \frac{2 \times 5}{(5^2 \times 2)^2}$   
㉠ = 10.

원본 이미지 다시보기

문제 선택하러 가기

# Streamlit Application

 **Re-TEST;**

 **복습할 문제를 선택해주세요**

여러 개의 문제 중 원하는 문제만 골라 시험지에 포함할 수 있어요.

\* problem\_1: ① ÷ ② = 1/250 입니다. ①에 알맞은 수 중에서 가장 작은 수는 얼마입니까?

선택 완료 및 시험지 생성

---

 [이미지 분석 화면으로 돌아가기](#)

문제 여러 개 중, 시험지로 구성하고 싶은 문제  
선택 후 시험지 markdown 형태 생성

 **Re-TEST;**

 **복습용 시험지 생성 중...**

선택한 문제들을 기반으로 시험지 형식으로 구성 중입니다.

복습용 시험지 생성 완료!

 시험지 출력

```
복습 시험

1. 쪽

1. ① ÷ ② = 1/250 입니다. ①에 알맞은 수 중에서 가장 작은 수는 얼마입니까?

2. 쪽

(빈칸)
```

 [해설지도 생성할까요?](#)

---

 [문제 선택 화면으로 돌아가기](#)

# Streamlit Application

## Re-TEST;

해설지 생성 중...

선택한 문제에 대한 자세한 해설을 구성 중입니다.

해설지 생성이 완료되었습니다!

해설지 출력

```
GCD(15, 40, 20) = 5. LCM(15, 40, 20) = 120. **가능**
20: GCD(15, 40, 25) = 5. LCM(15, 40, 25) = 600. **불가능**
30: GCD(15, 40, 30) = 5. LCM(15, 40, 30) = 120. **가능**
35: GCD(15, 40, 35) = 5. LCM(15, 40, 35) = 840. **불가능**
45: GCD(15, 40, 45) = 5. LCM(15, 40, 45) = 360. **불가능**
```

따라서, 가능한 ㉠은 5, 10, 20, 30 입니다.

서로 다른 세 자연수 조건 검토: 문제에서 서로 다른 세 자연수라고 했으므로, 15와 40이 아닌 수들만 고려해야 합니다. 이미 위에서 15와 40을 제외하고 검토했으므로 해당 조건은 만족합니다.

정답: 4개

이 될 수 있는 수는 5, 10, 20, 30 이므로 모두 4개입니다.

정답: 4개

시범지 화면으로 돌아가기

## Re-TEST

### 문제 1

서로 다른 세 자연수 15, 40, ㉠이 있습니다. 이 세 수의 최대공약수는 5이고 최소공배수는 120일 때, ㉠이 될 수 있는 수는 모두 몇 개입니까? (단, ㉠은 50보다 작은 수입니다.)

### 해설

1. 문제 이해:  
세 자연수 15, 40, ㉠의 최대공약수가 5이고 최소공배수가 120이라는 조건을 이용하여 ㉠이 될 수 있는 값을 찾고, 그 개수를 구하는 문제입니다.

2. 핵심 개념:  
• 최대공약수 (GCD): 주어진 수들의 공통 약수 중 가장 큰 수  
• 최소공배수 (LCM): 주어진 수들의 공통 배수 중 가장 작은 수  
• 소인수분해: 자연수를 소수들의 곱으로 나타내는 것

3. 풀이 과정:  
• 소인수분해:  
• 15 = 3 × 5  
• 40 = 2<sup>3</sup> × 5  
• 120 = 2<sup>3</sup> × 3 × 5  
• ㉠의 형태 추론:  
• 최대공약수가 5이므로 ㉠은 반드시 5를 인수로 가져야 합니다. 즉, ㉠ = 5 × a (a는 자연수) 형태입니다.  
• 최소공배수가 120 = 2<sup>3</sup> × 3 × 5 이므로, ㉠은 120의 약수여야 하며, 2<sup>3</sup>, 3, 5 이외의 다른 소인수를 가질 수 없습니다.  
• ㉠이 될 수 있는 수 찾기:  
㉠은 5를 인수로 가지고, 50보다 작아야 합니다. 따라서 ㉠은 5, 10, 15, 20, 25, 30, 35, 40, 45 중 하나입니다. 이 중 15와 40은 문제에서 이미 주어졌으므로 제외합니다. 남은 후보는 5, 10, 20, 25, 30, 35, 45 입니다. 이제 각 후보에 대해 최대공약수가 5, 최소공배수가 120이 되는지 확인해야 합니다.  
• 5: GCD(15, 40, 5) = 5. LCM(15, 40, 5) = 120. 가능  
• 10: GCD(15, 40, 10) = 5. LCM(15, 40, 10) = 120. 가능  
• 20: GCD(15, 40, 20) = 5. LCM(15, 40, 20) = 120. 가능  
• 25: GCD(15, 40, 25) = 5. LCM(15, 40, 25) = 600. 불가능  
• 30: GCD(15, 40, 30) = 5. LCM(15, 40, 30) = 120. 가능  
• 35: GCD(15, 40, 35) = 5. LCM(15, 40, 35) = 840. 불가능  
• 45: GCD(15, 40, 45) = 5. LCM(15, 40, 45) = 360. 불가능

따라서, 가능한 ㉠은 5, 10, 20, 30 입니다.

서로 다른 세 자연수 조건 검토: 문제에서 서로 다른 세 자연수라고 했으므로, 15와 40이 아닌 수들만 고려해야 합니다. 이미 위에서 15와 40을 제외하고 검토했으므로 해당 조건은 만족합니다.

4. 정답:  
㉠이 될 수 있는 수는 5, 10, 20, 30 이므로 모두 4개입니다.

문제지 / 해설지 PDF

생성 완료!

# 시연 영상 확인

Re-TEST; 2025 Capstone Design II

고려대(세종) 컴퓨터융합소프트웨어학과

2020270647 문창집

dwhouse0@korea.ac.kr



# 추가 개선 사항 / 한계점

## 🔧 개선 방향

Gemini 2.0 Flash 모델 API를 통해 이미지 입력, 텍스트 출력 수행

(기존 모델 Fine-Tuning 과정에서 kernel 충돌.. / OCR 성능 미흡)

API가 아닌 로컬에서 직접 수행할 수 있도록 모델을 경량화

다양한 데이터를 수집하여 모델의 성능을 더 높힐 것

## 🤖 향후 모델 구성 계획

1. 문제 영역 전용 Custom YOLOv8 Transfer Learning
2. OCR, 수식 인식, Donut(Naver Clova)이 아닌 TrOCR Fine-Tune
3. Gemma 3 27B Quantized 모델 Fine-Tune을 통한 경량화

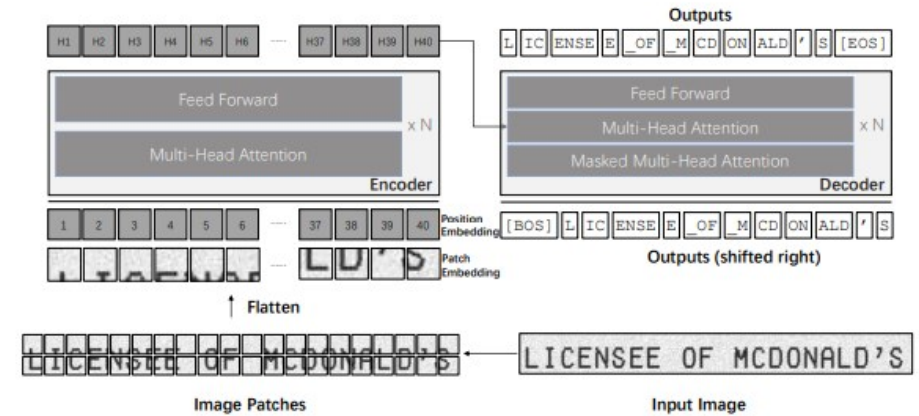


Figure 1: The architecture of TrOCR, where an encoder-decoder model is designed with a pre-trained image Transformer as the encoder and a pre-trained text Transformer as the decoder.

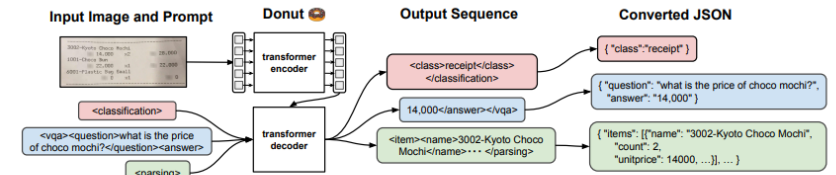


Fig. 3. The pipeline of Donut. The encoder maps a given document image into embeddings. With the encoded embeddings, the decoder generates a sequence of tokens that can be converted into a target type of information in a structured form

# 프로젝트를 통해 얻은 인사이트

Re-Test, 그리고 그 다음



학부 공부를 그동안 해오면서 잊고 있었던,  
“실생활에서 만난 문제점을 기술로 해결“ 하는 것의  
매력을 다시 떠올릴 수 있었던 경험.

UI/UX를 구성할 때 다양한 시나리오 설계  
사용자가 문제를 고르고, 결과물을 얻기까지의 흐름이  
명확하고 직관적이어야 한다!

내가 원하는 수준의 모델을 위해서는 정말 고품질의  
데이터가 필요하다는 것.

AI는 정말 무궁무진한 분야  
AI를 통해 사회문제를 해결할 수 있도록 더 연구할 것

# QNA?

## 2025 Capstone Design II

고려대(세종) 컴퓨터융합소프트웨어학과

2020270647 문창집

dwhouse0@korea.ac.kr



# 감사합니다.

Re-TEST; 2025 Capstone Design II

고려대(세종) 컴퓨터융합소프트웨어학과

2020270647 문창집

dwhouse0@korea.ac.kr

