



Києво-Могилянська академія

# РОЗРОБКА ІНСТРУМЕНТІВ ДЛЯ СТВОРЕННЯ БАЗИ НАВЧАЛЬНИХ МАТЕРІАЛІВ З МАТЕМАТИКИ ДЛЯ MATHPAR

Виконав: Поліщук В. Д.

Науковий керівник: д.ф.-м.н., проф. Малашонок Г. І.

19.05.2025

# Мета та завдання

**Мета:** Розробка системи для розпізнавання, структуризації та автоматичного розв'язання математичних задач з покроковими поясненнями

## Завдання:

- створити структурований набір даних з підручників
- розробити систему автоматичного розв'язання задач
- забезпечити інтеграцію з освітньою платформою

# Порівняльний аналіз методів розпізнавання

| Система     | CER (формули), % | FRA (сторінки), % | Українська мова, % | Швидкість (на 1 сторінку PDF) |
|-------------|------------------|-------------------|--------------------|-------------------------------|
| Mathpix     | 0,12             | 99                | 92                 | 1,2 сек                       |
| SimpleTex   | 1,3              | 68                | 0                  | 2,9 сек                       |
| InftyReader | 8,2              | 44                | 0                  | 7,5 сек                       |
| Nougat      | 10,3             | 10                | 0                  | 5,3 сек                       |

**Висновок:** Mathpix - єдина система з підтримкою української мови та найвищою точ

# Конвертація та структуризація PDF

Етапи обробки підручників:

- **Конвертація:** 20+ підручників → Mathpix API → текст+LaTeX
- **Парсинг:** Система на Python з регулярними виразами
- **Валідація:** Фільтрація некоректних задач
- **Класифікація:** Random Forest

# Набір даних

Створено структурований набір даних:

- **10,000+** математичних задач
- **20+** підручників для 1-7 класів
- **12** тематичних категорій

```
{
  "instruction": {
    "raw": "Розв'яжіть рівняння:\\\\ $x^{4}-4 x^{2}-5=0$.\\\\",
    "text": "Розв'яжіть рівняння:  $x^4-4 x^2-5=0$ .",
    "math_tokens": [
      "$x^{4}-4 x^{2}-5=0$"
    ]
  },
  "metadata": {
    "position": {
      "start_line": 1191,
      "end_line": 1191
    },
    "book": "kharkiv_2012_zbirnyk_dlya_abityrientiv",
    "problem_id": "10.10",
    "section": "Раціональні рівняння вищих степенів"
  }
}
```

# Система розв'язання

Модульна архітектура з патернами проектування:

- **Strategy** для різних типів розв'язувачів
- **Factory** для створення розв'язувачів
- **Composite** для математичних виразів

Розроблені модулі:

- **ArithmeticSolver**: PEMDAS, стовпчикові операції
- **EquationSolver**: лінійні рівняння з покроковими перетвореннями
- **Багатоформатна підтримка**: текст, LaTeX, JSON, MathPar



# Інтеграція з MathPartner

Система генерації навчальних матеріалів:

- Автоматична конвертація у формат для MathPartner
- Фільтрація за темами та типами задач

The screenshot displays four panels illustrating the conversion process:

- Top Left Panel:** Shows a user-friendly task: "TASK 05. Обчисліть  $756/7 = ?$ ". The output is "?".
- Top Middle Panel:** Shows the converted task: "TASK 10. Розв'яжіть рівняння  $724 \cdot x + 2 = 2$ ". The output is "2".
- Top Right Panel:** Shows the converted task: "TASK 01. UA-SM-Zaika02". The task is "Обчисліть  $368+46=?$ ". The output is "?".
- Bottom Left Panel:** Shows the detailed solution for the first task: "05. Розв'язок." followed by a long division of 756 by 7, resulting in 108. The final output is "756/7=108".
- Bottom Middle Panel:** Shows the detailed solution for the second task: "10. Розв'язок." followed by three steps: 1. "Початкове рівняння  $724 \cdot x + 2 = 2$ ", 2. "Перенесемо числа в праву частину  $724 \cdot x = 2 - 2 = 0$ ", 3. "Поділимо обидві частини на 724  $x = 0$ ". The final output is "0".
- Bottom Right Panel:** Shows the detailed solution for the third task: "01. Розв'язок." followed by the addition  $368 + 46 = 414$ . The final output is "414".

At the bottom of the interface, the numbers 108, 8, and 414 are displayed, corresponding to the results of the three tasks shown.

# Висновки

Створена система успішно структурувала понад 10 тисяч математичних задач з українських підручників та генерує для них покрокові розв'язки з поясненнями українською мовою. Система повністю інтегрована з освітньою платформою MathPartner і готова до практичного використання.

Рішення автоматизує створення навчальних матеріалів для вчителів та надає учням доступ до персоналізованих завдань. Створений датасет відкриває можливості для розробки ШІ-систем у математичній освіті.

**Дякую за увагу!**