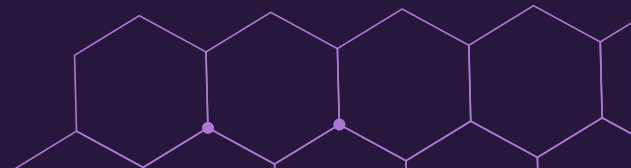
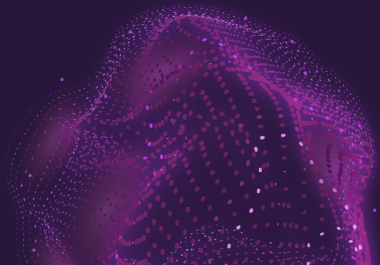




AI Meets Mathematics

An Interactive Chatbot for Primary Math



WORKFLOW



Majority of parents feel helpless with kids' homework – study

by Brett Henebery
30 Mar 2022

SHARE



The majority (67%) of parents worry they won't be able to effectively help their kids with homework this year, according to new research.

Free e-newsletter - get the latest news

LATEST NEWS

School's digital makeover sets learning standards

A leading Toowoomba girls school with an edtech giant is reaping the benefits

Australian kids need in good behaviour

A Senate inquiry into Australian schools?

Revealed: The 2022

THE STRAITS TIMES

SINGAPORE

LOG IN SUBSCRIBE

Parents go 'back to school' to learn how to coach their children in maths



Mr Muhammad Faizal Jamil and his wife Normaya Johari, with their daughter Nayli Yusairah and son Izz Haziq. ST PHOTO: GUYO ESP

Which of the following can react with an acid?

1. CO
2. Na_2CO_3
3. SO_2
4. O

Basic oxides: metal oxides
Acidic oxides: non-metal oxides
Amphoteric oxides: CO

PROBLEM STATEMENT

×

×

- Our AI-powered chatbot addresses the need for **timely academic support**, offering **instant and personalized assistance** to upper primary students. It fills the gaps in traditional learning by providing on-demand explanations, solutions, and concept clarifications, enhancing student engagement and understanding in mathematics
- Success metrics: RAGAS score, Answer Relevancy & Faithfulness





DATA

PRIMARY MATH MCQS



- ◆ **80** mathematics multiple choice questions taken from past year PSLE papers
- ◆ Restricted to questions without pictures

Find the remainder of 3908 divided by 7.

- | | |
|-------|-------|
| (1) 2 | (3) 5 |
| (2) 3 | (4) 6 |

What is five million, nine thousand and sixty in numerals?

- (1) 5 009 006
- (2) 5 009 060
- (3) 5 090 006
- (4) 5 090 060



FEATURE ENGINEERING (OPENAI)

SOLUTION

```
prompt = (  
    "You are an AI capable of solving math problems. "  
    "For each problem, select the best answer from the given multiple-choice  
    options. "  
    "Explain why this choice is correct and why the others are not. "  
    "Your explanations should be clear and concise, using simple arithmetic and  
    intuitive reasoning when possible. "  
    f"Here is the problem: {question}\n\n"  
    "Best Answer and Explanation: \n\n"  
)
```

Express 1.8% as a decimal. (1) 0.0018 (2) 0.018 (3) 0.18 (4) 18

The best answer is (2) 0.018.

Explanation: To convert a percentage to a decimal, you divide by 100. So, 1.8 divided by 100 equals 0.018. The other options are incorrect because they do not represent the correct decimal equivalent of 1.8%. For example, 0.0018 is equivalent to 0.18%, 0.18 is equivalent to 18%, and 18 is equivalent to 1800%, not 1.8%.

COGNITIVE LEVEL

```
prompt = (  
    "Here are the classifications:\n"  
    "A01: Recall mathematical facts, concepts, rules, and formulae; perform  
    straightforward computations and algebraic procedures.\n"  
    "A02: Interpret information; understand and apply mathematical concepts and  
    skills in a variety of contexts.\n"  
    "A03: Reason mathematically; analyze information and make inferences; select  
    appropriate strategies to solve problems.\n\n"  
    f"Classify the following math question into A01, A02, or A03:\n\nQuestion:  
{question}\n\nClassification:"  
)
```

Express 1.8% as a decimal. (1) 0.0018 (2) 0.018 (3) 0.18 (4) 18
Classification: A01

TOPIC

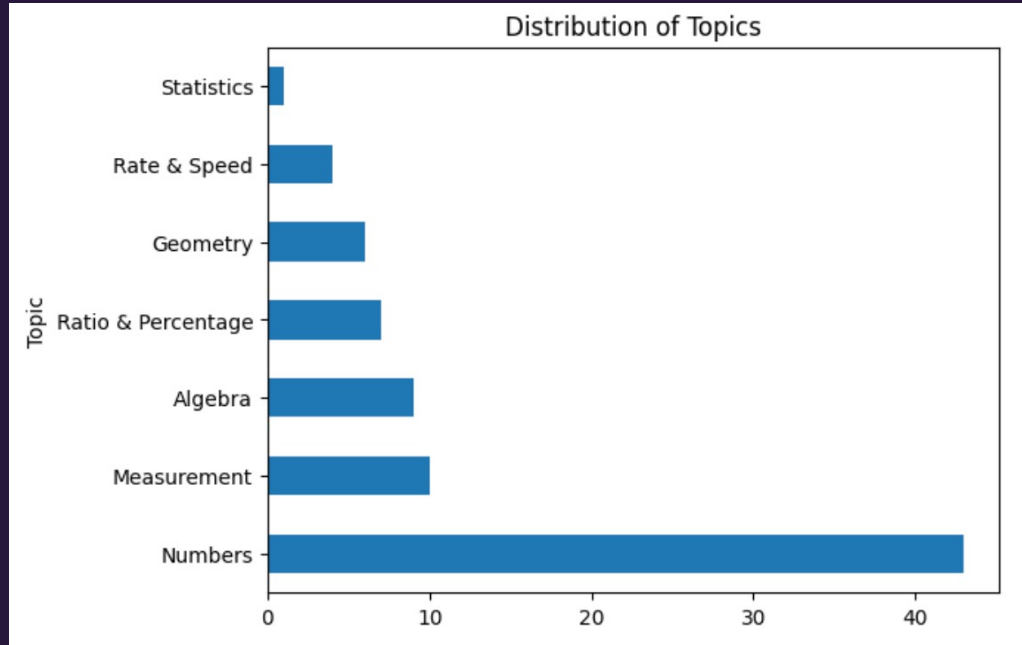
```
prompt = (  
    "Here are the topics:\n"  
    "Numbers, Measurement, Statistics, Geometry, Ratio & Percentage, Rate & Speed,  
    Algebra"  
    f"Classify the following math question into one of the relevant  
    topics:\n\nQuestion: {question}\n\nTopic:"  
)
```

Express 1.8% as a decimal. (1) 0.0018 (2) 0.018 (3) 0.18 (4) 18
Ratio & Percentage

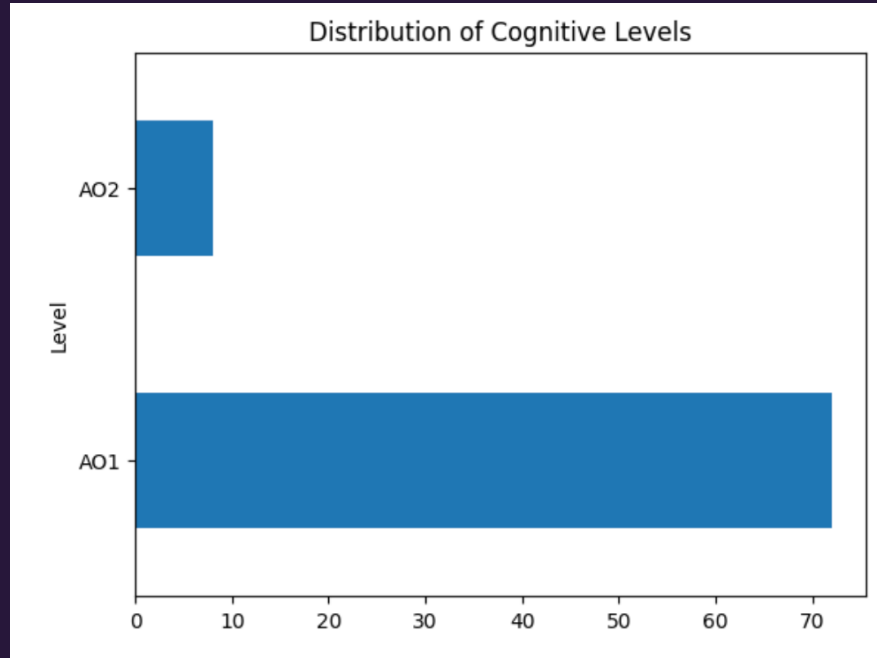


DATA EXPLORATION

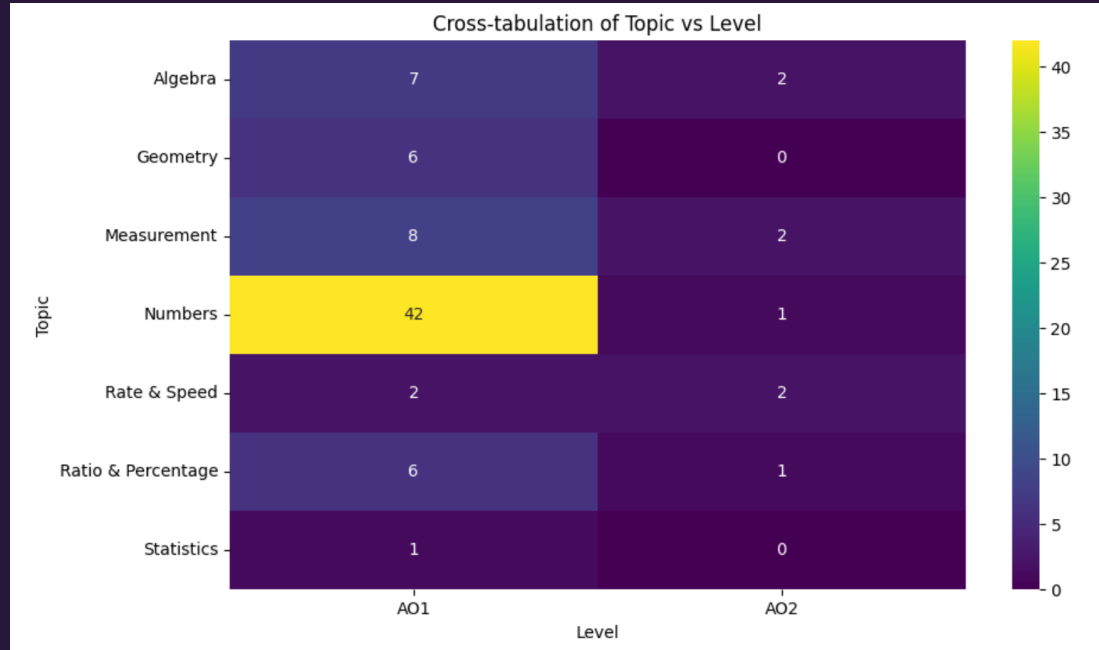
TOPIC DISTRIBUTION



COGNITIVE LEVEL DISTRIBUTION



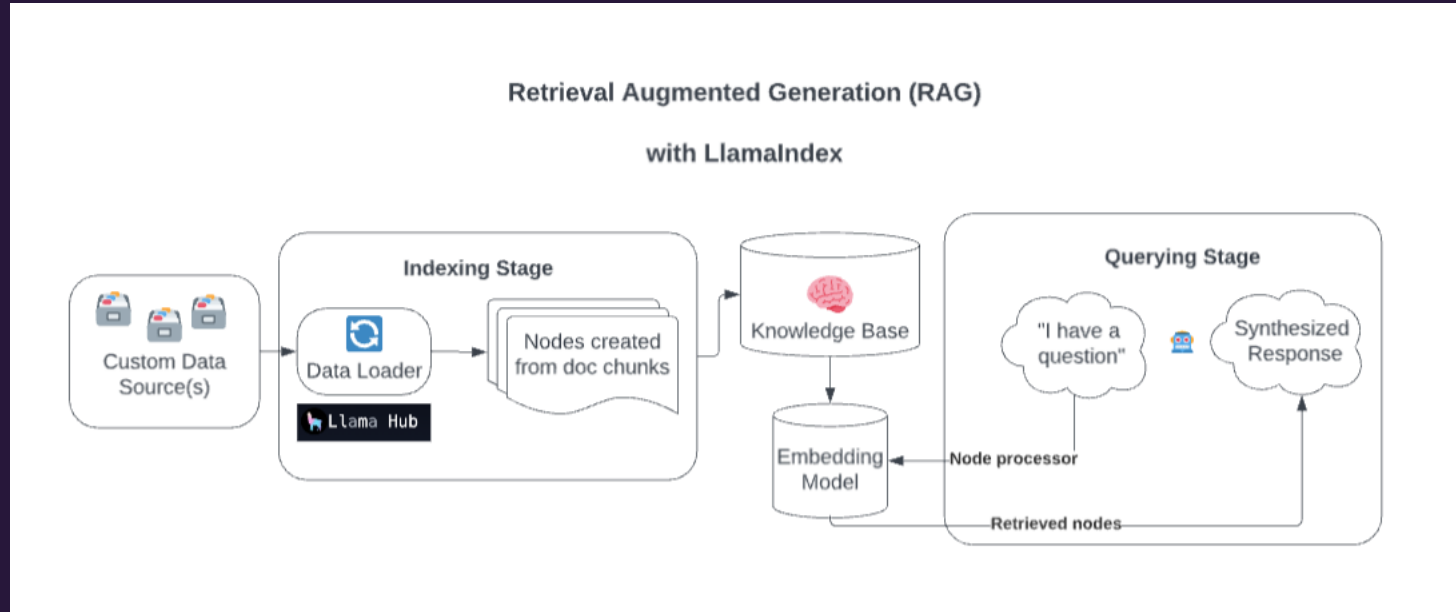
TOPIC VS COGNITIVE LEVEL





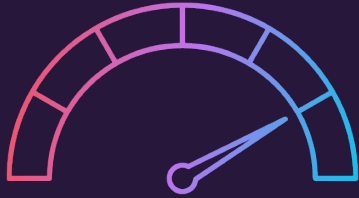
BUILDING THE CHATBOT

BIG PICTURE



Source: <https://discuss.streamlit.io/t/build-a-chatbot-with-custom-data-sources-powered-by-llamaindex/49973>

METRICS



RAGAS score
↑ better



Answer Relevancy
↑ better



Faithfulness
↑ better

MODELING



	RAGAS Score	Answer Relevancy	Faithfulness
GPT 3.5 Turbo	0.7685	0.8517	0.7000
Finetuned	0.8150	0.8922	0.7500

COMPARING DIFFERENCES

700 305 is _____ more than 680 305.

A) 200 tens B) 2 thousands C) 20 thousands (correct) D) 2 ten thousands.

GPT 3.5 TURBO

The answer is B) 2 thousands.

FINE TUNED

The correct answer is C) 20 thousands.

COMPARING DIFFERENCES

What is the value of 3 in 3560?

A) 3 thousands B) 3 hundreds C) 3 tens D) 3 ones

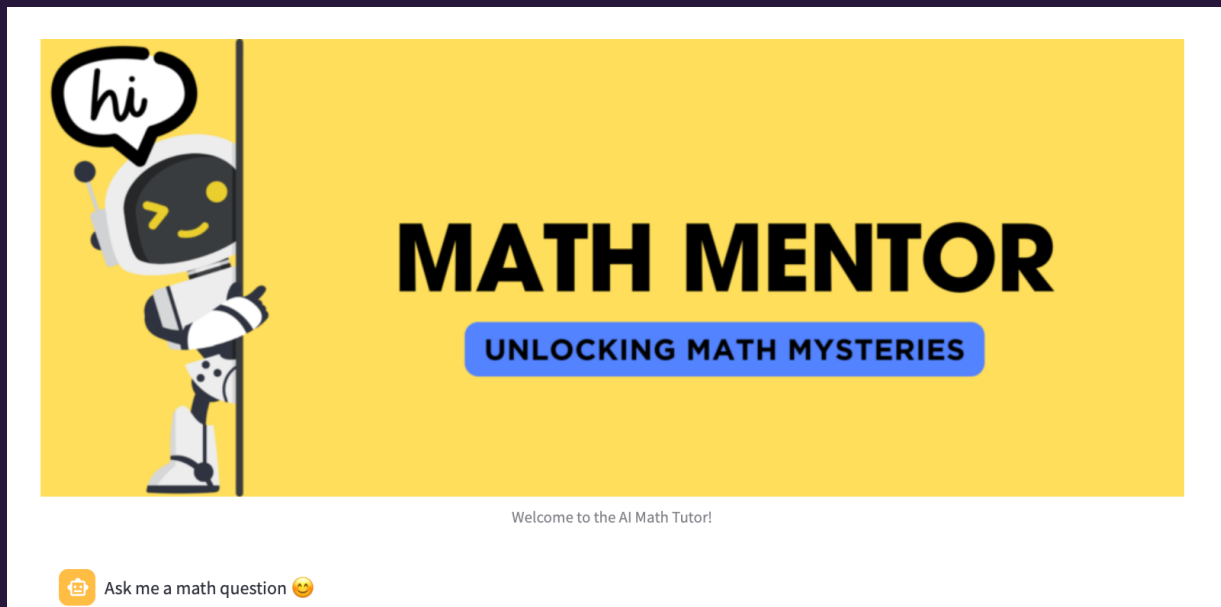
GPT 3.5 TURBO

3 hundreds

FINE TUNED

The value of 3 in 3560 is 3 thousands.

STREAMLIT DEMONSTRATION



REFERENCE

- <https://www.straitstimes.com/singapore/education/families-spent-14b-on-private-tuition-for-kids-last-year-as-parents-fork-out>
- <https://www.straitstimes.com/singapore/education/7-in-10-parents-send-their-children-for-tuition-st-poll>
- <https://www.straitstimes.com/singapore/education/parents-go-back-to-school-to-learn-how-to-coach-their-children-in-maths>
- <https://www.theeducatoronline.com/k12/news/majority-of-parents-feel-helpless-with-kids-homework--study/279916>