**3rd Place Solution (with Magic Boost)**

First, I want to thank the organizers for hosting this competition, and the Kaggle community for all the discussions and shared notebooks that helped me learn so much.

**My Solution**

A two-stage approach:

**Stage 1: Retrieval**

* Used Qwen-14B Embedder (from [@anhvth226](https://www.kaggle.com/anhvth226) ) ensemble with Qwen-14B embedder trained with FlagEmbedding
* Retrieved 35 most relevant misconceptions for each question
* Used publicly shared Retriever and tried my own FlagEmbedding-trained Retriever
* To be honest, my self-trained Retriever was pretty bad and removing it actually improved the score (lol)

**Stage 2: Reranking**

* Used Qwen-32B-instruct-AWQ reranker
* Ensembled 6 different LoRAs with various training parameters and cross-validation
* One model included about 2000 GPT4-mini generated samples, but the improvement was… not very significant

Scores at this point:

* Public LB: 0.590
* Private LB: 0.564

**Time for Some Magic!**

Here comes the most interesting part!

It started when I discovered over 900 unseen misconceptions in the misconceptions table that never appeared in the training data. Plus, according to [@zhudong1949](https://www.kaggle.com/zhudong1949) and [@eugenkrylov](https://www.kaggle.com/eugenkrylov) findings (check here: <https://www.kaggle.com/competitions/eedi-mining-misconceptions-in-mathematics/discussion/550875)>, the testing set had many unseen subjects and only about 685 questions.

So I thought: these unseen misconceptions must make up a significant portion of the testing set!

**Let's Experiment**

On the second-to-last day of the competition (yes, I really dared to use two submissions for experiments at this point XD), I did this LB probing:

* Predicted only 1 misconception per question
* Tested twice:
  1. Using only seen misconceptions from training data: got 0.154
  2. Using only unseen misconceptions: got 0.444

Seeing these results, I made a bold guess that the seen-to-unseen ratio in the testing data was roughly 1:3.

**The Final Magic Touch**

So I implemented a simple post-processing:

* Multiplied the probabilities of all predicted unseen misconceptions by a constant C
* Adjusted until unseen misconceptions made up 75% of the first predictions

This little magic trick made the scores skyrocket 😮:

* Public LB: 0.590 → 0.658
* Private LB: 0.564 → 0.600

In the final version, I also randomly shuffled the order of misconceptions input to the reranker, which gave a small additional boost:

* Public LB: 0.670
* Private LB: 0.602

**Code**

The inference code can be found at:  
  
<https://www.kaggle.com/code/threerabbits/eedi-11-21-myq14b-q32b-rerank-mod-novel-local-suf>