Automatic Detection and Association of Misconceptions in Mathematics Using LLMs

Navarro Steven, Amaya Juan David, Garcia David Santiago
District University Francisco Jose de Caldas, Colombia, Bogotá

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Introduction

Understanding students' misconceptions in math is key to improving education. While this has traditionally required manual expert analysis, AI is now used to automate the process. However, challenges remain in achieving accurate, relevant, and scalable results in real-world educational settings.

Goal

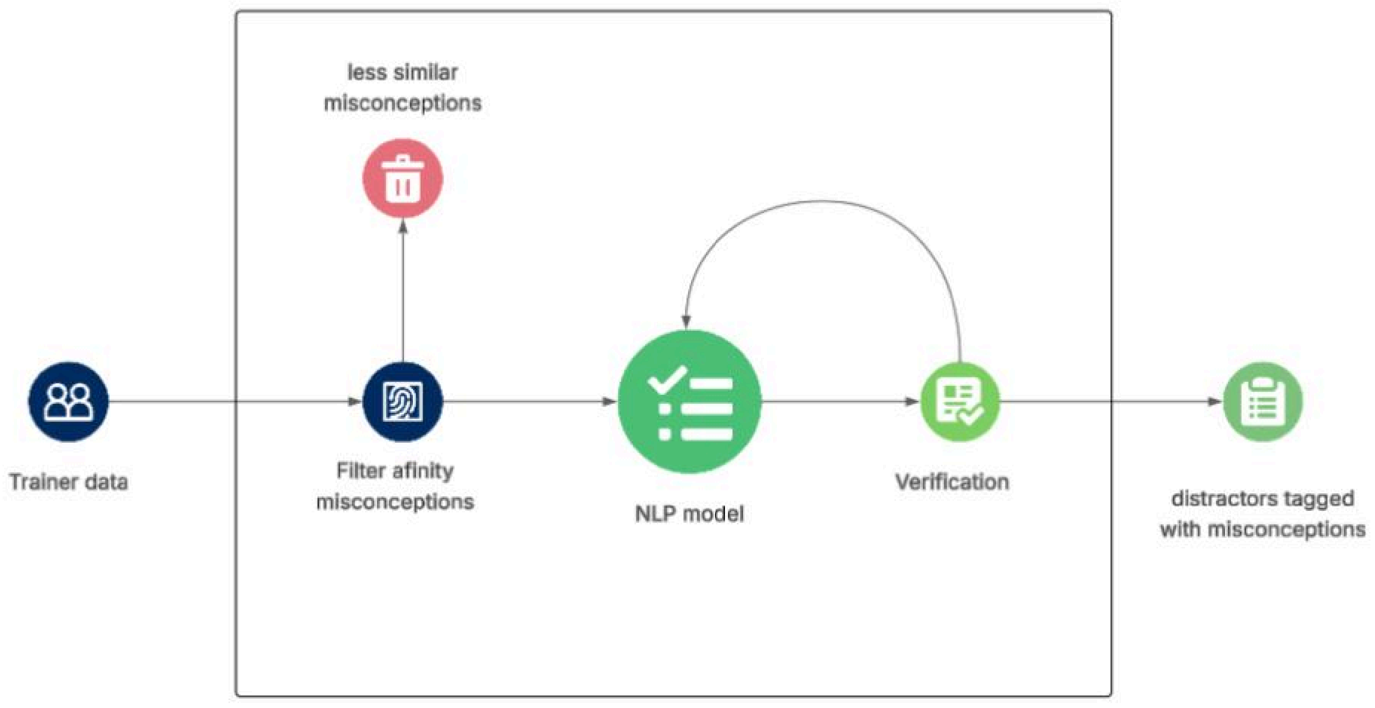
This research explores how large language models can be used to accurately detect and link misconceptions to distractors in math assessments. The goal is to develop a system that can automatically label at least 25 misconceptions per distractor with high semantic relevance.

Proposed Solution

Our proposed system is a modular, AI-powered tool that detects and associates misconceptions with distractors in mathematics questions using large language models (LLMs). The solution follows a structured pipeline:

- **Data Preparation:** Questions, correct answers, distractors, and a misconception bank are preprocessed and fed into the system.
- **Filtering:** An initial filter discards low-relevance misconceptions using affinity thresholds.
- **NLP Analysis:** Claude 3.5 Sonnet is employed to associate misconceptions with distractors based on textual and conceptual similarities.
- **Verification and Feedback:** GPT-4 validates the associations; if the confidence is low, the model is retrained or fine-tuned.

This iterative approach ensures pedagogical relevance, technical accuracy, and adaptability to diverse datasets.



Experiments

We tested our system with the EEDI Kaggle dataset using a local LLM to label misconceptions based on semantic similarity between questions and distractors. Although slower, the model runs on limited resources, offering a practical and scalable alternative to larger models that require paid APIs.

Results

The system analyzed over 4,000 question–distractor pairs from the EEDI dataset, filtering them by semantic affinity. A local LLM was used to generate misconception labels with consistent accuracy across various math topics such as geometry and algebra. While the model was slower than cloud-based alternatives, it ran reliably on limited hardware and produced meaningful predictions aligned with real misconceptions.

Conclusions

This project demonstrated that a local LLM can effectively detect and label mathematical misconceptions in multiple-choice assessments. Despite its slower processing time, the model offered strong semantic understanding without relying on external APIs. Its offline, resource-efficient design makes it a practical solution for educational environments with limited computational access.

Bibliography

Eedi - Mining Misconceptions in Mathematics. (s. f.). Kaggle. https://www.kaggle.com/competitions/eedi-mining-misconceptions-in-mathematics/overview