# Evaluation and Analysis

## Project Description

The primary goal of this project is to leverage the latest advancements in HMER and MLP to develop a comprehensive solution for the automatic grading of handwritten math assignments at the school level. By combining these technologies, this project aims to address the challenges associated with grading handwritten math assignments and enhance the efficiency and effectiveness of math education.

## Evaluation Metrics

The evaluation metrics employed in CROHME, particularly in the 2023 edition, focus on symbol-level evaluation using the Symbol Layout Graph (SymLG). This representation allows for direct comparison among systems regardless of whether they produce stroke-level or symbol-level results. The primary metrics used include:

- Expression recognition rate: Percentage of expressions with matching MathML trees and at most 3 incorrect symbols or relations.

- Structure recognition rate: Percentage of expressions with completely correct recognition, considering both symbols and their relations.

## Criteria for Evaluating Research Finding

In evaluating research findings using the CROHME dataset, the following criteria are considered:

- Accuracy: The accuracy of recognizing handwritten mathematical expressions, both in terms of individual symbols and the overall structure, is of paramount importance.

- Robustness: Systems should demonstrate robustness across different input formats and handwriting styles to be considered effective.

- Efficiency: The efficiency of the recognition system, including computational complexity and runtime, is crucial for practical deployment.

## Interpretating Results

Interpreting the results obtained from the evaluation involves considering both the expression recognition rate and structure recognition rate. A higher expression recognition rate indicates the ability of the system to recognize symbols and their relations to a certain degree of accuracy. On the other hand, a higher structure recognition rate signifies the system's capability to accurately reconstruct the entire expression, including its structural hierarchy.

In cases where the structure recognition rate is significantly lower than the expression recognition rate, it indicates that the system struggles with understanding the hierarchical relationships between symbols, leading to structural errors. Conversely, if the structure recognition rate is high but the expression recognition rate is low, it suggests that the system may correctly identify symbols but fails to recognize their relationships accurately.

## Improving Performance

To improve the performance of formula recognition systems using the CROHME dataset, several strategies can be employed:

- Enhance Symbol Recognition: Improving the accuracy of individual symbol recognition is crucial for achieving higher expression recognition rates. This can be accomplished through advanced feature extraction techniques and more sophisticated classification algorithms.

- Refine Structural Understanding: Addressing the challenges related to understanding the hierarchical structure of mathematical expressions is essential. Utilizing deep learning architectures capable of capturing long-range dependencies may help improve structure recognition rates.

- Data Augmentation and Transfer Learning: Augmenting the dataset with diverse handwriting styles and leveraging transfer learning techniques from related tasks could enhance the robustness of recognition systems.