Automated Grading of Handwritten Assignments in Mathematics

—

Sumukhi Ganesan
DS 5500 - Data Science Capstone Project

Abstract / Agenda

Objective: Automating the grading of handwritten math assignments to enhance school-level math education.

Problem: Grading handwritten math assignments is challenging due to the combination of math expressions and natural language.

Approach: Integrating Handwritten Mathematical Expression Recognition (HMER) with Mathematical Language Processing (MLP) for automatic grading.

Methodologies: Utilizing the Dense-WAP method for HMER and MLP-S, MLP-B algorithms for evaluation.

Expected Benefits: Simplified submission process for students, reduced grading time for teachers, and improved learning experience.

Anticipated Results: High accuracy in structure recognition with Dense-WAP method, promising outcomes from MLP-S and MLP-B algorithms.

Target Audience: Middle school and high school educators, teachers, and students seeking to optimize math assessment processes.

Introduction

Motivation: Traditional methods of grading handwritten math assignments in schools are time-consuming and often subject to human error. With the increasing emphasis on personalized learning and efficient assessment practices, there is a growing need for automated grading solutions in math education.

Problem: Grading math assignments presents a unique challenge due to the combination of mathematical expressions and natural language. Existing grading systems are often ill-equipped to handle this complexity, resulting in inefficiencies and inconsistencies in the assessment process. Handwritten responses add another layer of complexity in the form of recognizing and interpreting math expressions.

Objective

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.

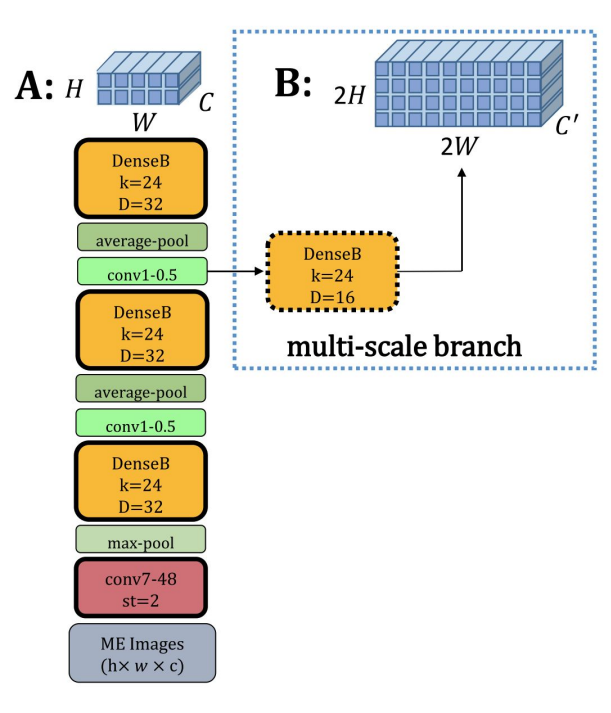
Literature Review

The literature surrounding automatic grading of math assignments and Handwritten Mathematical Expression Recognition (HMER) is extensive, with numerous papers exploring these topics independently. However, there is a notable gap in the literature regarding the integration of these two areas. While there are several papers focusing on auto-grading methods and others addressing HMER techniques, few studies have investigated the combined approach of leveraging HMER alongside mathematical language processing for automated grading of handwritten math assignments. This gap in research highlights the potential for innovation and advancement in educational technology by integrating these complementary fields.

Methodology - HMER

Zhang, J., Du, J., and Dai, L. introduced the Dense-WAP method for Handwritten Mathematical Expression Recognition (HMER) in their paper titled "*Multi-Scale Attention with Dense Encoder for Handwritten Mathematical Expression Recognition*," presented at the International Conference on Pattern Recognition (ICPR) in 2018.

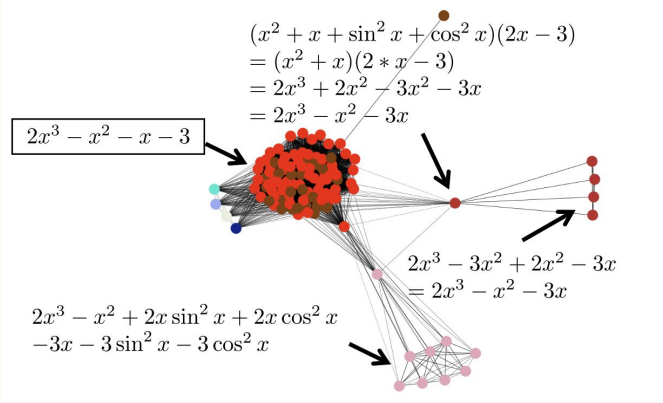
The attention based encoder-decoder model recognizes mathematical expression images from two-dimensional layouts to one-dimensional LaTeX strings. Zhang et al., improved the encoder by employing densely connected convolutional networks as they can strengthen feature extraction and facilitate gradient propagation especially on a small training set.



Methodology - Grading

Andrew S. Lan, Divyanshu Vats, Andrew E. Waters, and Richard G. Baraniuk presented their research on Mathematical Language Processing (MLP) in the paper "*Mathematical Language Processing: Automatic Grading and Feedback for Open Response Mathematical Questions*" at the ACM Conference on Learning @ Scale in 2015.

The Mathematical Language Processing (MLP) algorithm aims to address the challenge of automatically grading open-response mathematical questions common in STEM courses. Leveraging solution data from a large number of learners, MLP employs a data-driven framework to evaluate correctness, assign partial-credit scores, and provide feedback. Inspired by natural language processing techniques, MLP involves converting solutions into numerical features, clustering these features to identify correct, partially correct, and incorrect solutions, and then automatically grading remaining solutions based on their assigned cluster.

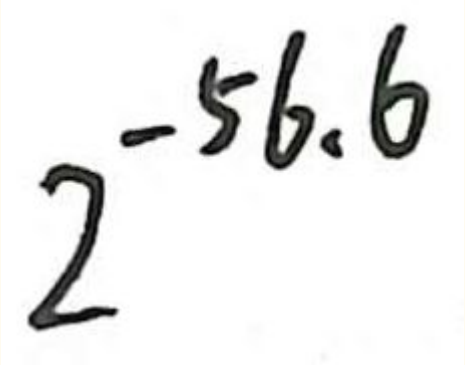


Dataset

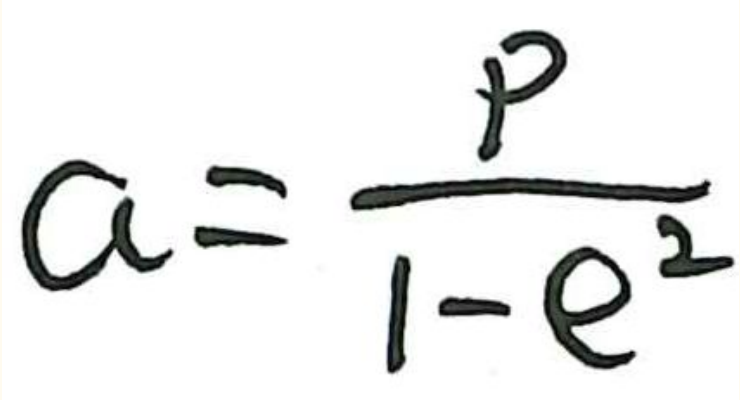
Xie, Y. et al. introduced the ICDAR 2023 CROHME dataset for the Competition on Recognition of Handwritten Mathematical Expressions. This dataset, presented in the proceedings of the International Conference on Document Analysis and Recognition (ICDAR) 2023, provides a valuable resource for evaluating and benchmarking HMER algorithms.

Tasks	Train	#	Validation	#	Test	#
Task 1						
On-line	Train 2019	9 993	Test 2016	1147		
	Validation 2019	986				
	artificial	145 108				
	new samples	1045	new samples	555	new samples	2 300
	Total =	**157 132**	**Total =**	**1 702**	**Total =**	**2 300**
Task 2						
Off-line	rendered	9975	rendered	1147		
	real	1604				
	OffRaSHME	10 000				
	new samples	1045	new samples	555	new samples	2 300
	Total =	**20 979**	**Total =**	**1 702**	**Total =**	**2 300**
Task 3:						
Bimodal	InkML + rendered	9 975	InkML + rendered	1147		
	InkML + real	1 604				
	new samples	1045	new samples	555	new samples	2 300
	Total =	**10 979**	**Total =**	**1 702**	**Total =**	**2 300**

Sample Images and Labels



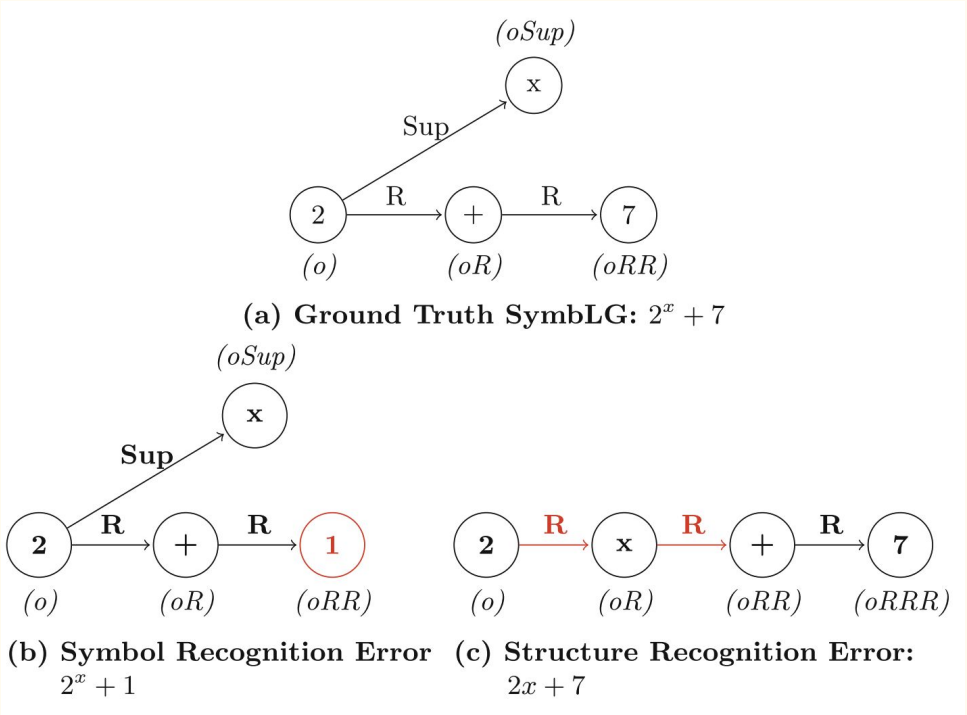
$2 ^ { - 5 6 . 6 }$



$a = \frac { p } { 1 - e ^ { 2 } }$

Evaluation

The CROHME dataset provides Symbol Layout Graphs (SymLGs) that can be used to evaluate HMER models. The evaluation metric of formula recognition is always the same symbol-level evaluation. Expression and Structure recognition rates are considered as important metrics.



(a) Ground Truth SymbLG: $2^x + 7$

(b) Symbol Recognition Error $2^x + 1$

(c) Structure Recognition Error: $2x + 7$

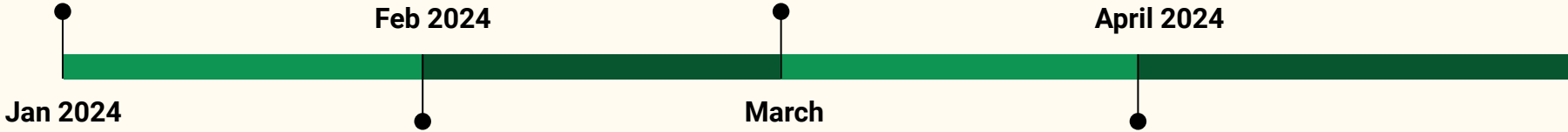
Project Plan

Proposal

Defining problem statement, conducting literature review, submitting project proposal

MLP Auto-grading

Test auto-grading based on MLP algorithms using collected responses, system integration and testing



HMER and Grading Data Collection

Complete handwritten response recognition using CROHME dataset, select questions for autograding and collect responses

Final Report and Presentation

Summarize results and draft the final project report, prepare and present poster

Conclusion

In conclusion, the proposed integration of Handwritten Mathematical Expression Recognition (HMER) and Mathematical Language Processing (MLP) presents a promising solution for automating the grading of handwritten math assignments at the school level. By leveraging advanced algorithms and methodologies, this project aims to streamline the assessment process, improve grading accuracy, and enhance the overall learning experience for students and educators. Through rigorous testing and validation, this project seeks to contribute to the advancement of educational technology and pedagogy in mathematics education.

References

https://github.com/sumukhig/ds5500-capstone-project/blob/main/docs/Literature%20Survey.docx