A Survey on AI-Based Answer Sheet Evaluation and Result Management Systems

Diya Kubal[1], Vaidehi Kokare[2], Aayush Dhotre[3], Dr. Riyazahemed Jamadar[4]

AISSMS Institute of Information Technology, Pune, India

kubaldiya6@gmail.com, kokarevaidehi2@gmail.com

dhotreaayush123@gmail.com, riyaz.jamadar@aissmsioit.org

Abstract: *Artificial Intelligence (AI) in teaching can transform traditional methods of assessment, particularly the methods of evaluation of the descriptive and open-ended responses of the students. Large language models (LLMs) can read, interpret and generate text in a human-like way, which means that automated grading systems can be both more scalable and objective than in the past when it was performed by human evaluators. Beyond accuracy, transparent, reliable and capable of producing actionable feedback post in accordance with pedagogical standards, many LLMs offer a black box design, which can actively raise concerns about fairness, interpretability and transparency, and the cost/benefit balance of processing speed and analytical depth can limit learners and institutions in their use.*

An emerging solution, which is the focus of a review in the present paper, is dual-process LLM frameworks of automated assessment. These hybrid systems are a combination of a bigger, more complex system that would perform more comprehensive rubric-based assessment and a system that would make detailed feedback with a slightly small model which could be used to score first assessments fast. To provide researchers and educators with a comprehensive outline of the development of reliable and efficient AI-powered assessment devices.

Keywords: *Automated Grading, Answer Sheet Evaluation, Natural Language Processing, Large Language Models, Dual-Process Framework, Education Technology*