

Fundamentos para la retroalimentación de evaluación formativa asistida por IA para tareas de respuesta corta en clases de inscripción grande

[Google traduccion del Inglés]

Matthew Beckman
Penn State University

13 Septiembre, 2022

Resumen

La investigación sugiere que las tareas de “escribir para aprender” mejoran los resultados del aprendizaje, pero los métodos de evaluación formativa de respuesta construida se vuelven difíciles de manejar con clases grandes. Este estudio evalúa algoritmos de procesamiento de lenguaje natural para ayudar a este objetivo. Seis tareas de respuesta corta completadas por 1935 estudiantes fueron calificadas por varios evaluadores humanos utilizando una rúbrica detallada y un algoritmo. Los resultados indican un acuerdo sustancial entre evaluadores utilizando kappa ponderado cuadrático para pares de evaluadores (cada $QWK > 0,74$) y consenso grupal ($Kappa \text{ de Fleiss} = 0,68$). Además, se estimó el acuerdo entre evaluadores para un evaluador que había obtenido 178 respuestas siete años antes ($QWK = 0,88$). Con un acuerdo convincente de los evaluadores, el estudio luego dirige el análisis de grupos del texto de respuesta para permitir a los instructores atribuir significado a los grupos como un medio para la evaluación formativa escalable.