

autogradeR: reproducibilidad desde la instrucción de ciencia de datos

Anónimo

Abstract Al abordar el problema apremiante de la falta de reproducibilidad en ciencia, autogradeR plantea revolucionar la enseñanza e investigación en ciencia de datos. autogradeR permitirá a estudiantes y profesores integrar prácticas de reproducibilidad en sus tareas y proyectos, promoviendo la generación de resultados consistentes y confiables. autogradeR estandariza respuestas a preguntas de naturaleza cerradas, lo que ahorrará tiempo a los instructores y les permitirá centrarse en aspectos más relevantes y enriquecedores de la enseñanza. autogradeR sentará las bases para una educación en programación que fomenta la investigación reproducible y de calidad, promoviendo resultados de aprendizaje rigurosos y un ambiente inclusivo en la ciencia de datos.

Palabras clave: Autograder - Aprendizaje automatico - Ciencia de datos - Pedagogia

La falta de reproducibilidad en más del 60 % de la investigación representa un obstáculo significativo para el avance en el conocimiento científico. Este patrón de falta de reproducibilidad termina socavando la confiabilidad y credibilidad de los hallazgos y la capacidad a futuro de construir conocimiento basado en experiencias previas. Estudios anteriores han reconocido la importancia y el impacto que la falta de reproducibilidad tiene sobre la adquisición de conocimiento en diferentes campos. Sin embargo, con frecuencia las discusiones sobre cambios que se pudiesen implementar para incrementar la reproducibilidad en contextos científicos están directamente relacionados a procesos investigativos. Por lo tanto, históricamente se ha dejado a un lado la reproducibilidad como un componente importante de la enseñanza en ciencia de datos. Este patrón donde la instrucción formal sobre reproducibilidad es secundario o inexistente es también replicado en otros campos que de una u otra forma las herramientas computacionales juegan un papel importante en el análisis de datos.

Reconociendo el problema apremiante de la reproducibilidad en contextos científicos, el paquete de R autogradeR se presenta como una iniciativa pionera con un potencial excepcional para reformar las metodologías de investigación científica y pedagógicas. El paquete autogradeR es en principio un paquete que permite la revisión y contraste rápido de información por parte de estudiantes e instructores en temas particulares donde los tipos de respuestas son limitados. Sin embargo, autogradeR tiene dos ventajas principales. Primero, autogradeR permite que estudiantes sean consistentes con las estructuras de respuesta y la importancia de generar tareas reproducibles por ellos mismos en otros computadores, sus compañeros e instructores. Segundo, al estandarizar aspectos que son de limitada envergadura en sus potenciales respuestas, autogradeR permite que instructores inviertan más tiempo preparando y discutiendo aspectos que van más allá de proveer una respuesta puntual. Por lo tanto, autogradeR se plantea en principio como un cambio en las prioridades de instrucción en cursos de ciencias de datos. El tiempo que se invierte en revisar cíclicamente respuestas cerradas puede ser invertido en permitir discusiones sobre el uso apropiado, la ética, entre otros, de prácticas asociadas a la ciencia de datos.

Al defender la ciencia abierta y reproducible, autogradeR sienta las bases para una nueva era de educación basada en programación, fomentando a largo plazo la investigación científica reproducible y de calidad. autogradeR tiene la capacidad de transformar no solo la educación en ciencia de datos, sino también allanar el camino hacia un entorno de enseñanza inclusivo, fomentando resultados que son consecuencia directa de aprendizaje rigurosos y con impacto.