

Aprendizaje estadístico aplicado para potenciar la enseñanza de inglés en primaria: El caso de Ceibal en Inglés en Uruguay

Palabras clave: Educational data mining, Aprendizaje Estadístico, tidymodels

Resumen

1. Introducción

Ceibal es el centro de innovación educativa con tecnologías digitales de Uruguay, cuya misión es promover la integración de la tecnología en la educación para mejorar los aprendizajes e impulsar procesos de innovación, inclusión y crecimiento personal. Ceibal en Inglés (CEI) es el programa de Ceibal para la enseñanza de inglés en centros de enseñanza pública de Primaria y Media (Secundaria y UTU). En el año 2021 se comienza a utilizar la plataforma Little Bridge donde los escolares realizan actividades asignadas por docentes remotos. En este trabajo, el objetivo global es estudiar cómo el uso de la plataforma LB está vinculado al desempeño académico de los niños en el aprendizaje del idioma inglés.

2. Los datos e indicadores de uso de plataformas.

Se trabajó con la población de estudiantes que abarca desde cuarto hasta sexto grado de educación primaria y que utilizó la plataforma Little Bridge. Se combinaron distintos conjuntos de datos, incluyendo variables sociodemográficas de los alumnos, actividades de alumnos y actividades de docentes en el Learning Management System (LMS), y el desempeño de los estudiantes en pruebas estandarizadas nacionales. A partir de los datos del LMS se desarrollaron indicadores relacionados con la motivación, comunicación, velocidad y rendimiento del estudiante. Debido al volumen de la información utilizamos la biblioteca `data.table` para su manipulación y limpieza.

3. Predecir desempeño

Utilizando el resultado de la prueba adaptativa en inglés como variable de respuesta y los indicadores de uso de LB en sus distintas dimensiones como variables explicativas, se ajustaron métodos de aprendizaje estadístico para predecir el desempeño de los estudiantes en pruebas estandarizadas nacionales. Para la exploración de modelos posibles nos apoyamos en las bibliotecas `tidymodels` y `h2o`. El objetivo fue detectar distintos comportamientos que permitan identificar a estudiantes que no alcancen un nivel suficiente de inglés en la prueba. Esta investigación es relevante para definir políticas educativas basadas en evidencia. Un desafío importante fue cómo lidiar con la estructura de datos de subgrupos en algoritmos de aprendizaje automático, generalmente desarrollados para observaciones independientes.