

## Predicción temprana de la deserción universitaria en primer año: Un modelo multivariable basado en aprendizaje automático.

Pablo Vilaboa, Roxana Martínez  
Universidad Abierta Interamericana (UAI)  
Ciudad Autónoma de Buenos Aires, Argentina  
{Pablo.Vilaboa,Roxana.Martinez}@uai.edu.ar

se orienta al estudio de las diferentes propuestas con el propósito de construir un marco estructurado que permita integrar dominios de datos contextuales (sociales, académicos, psicológicos e institucionales) del estudiante para establecer un mecanismo replicable de análisis de estos datos y traducir los resultados en lineamientos operativos para la toma de decisiones

### RESUMEN

La deserción estudiantil en los primeros años de la formación universitaria constituye un desafío significativo tanto para las instituciones educativas como para los propios estudiantes, dado su impacto directo en el aprovechamiento de los recursos institucionales y en la planificación de estrategias académicas orientadas a cada cohorte.

En este contexto, la disponibilidad de herramientas que permitan identificar de manera temprana a los estudiantes en riesgo resulta fundamental para optimizar la asignación de recursos y fortalecer las políticas de acompañamiento académico. La incorporación de modelos predictivos basados en múltiples variables posibilita una detección más precisa y oportuna de situaciones de riesgo, superando las limitaciones de los enfoques tradicionales. Asimismo, el desarrollo de un modelo predictivo multivariable complementado con una metodología que oriente su implementación en el ámbito institucional representa un enfoque innovador; al integrar información proveniente de distintos dominios y contribuir tanto al avance del conocimiento académico como a la mejora de los sistemas de alerta temprana.

La investigación se fundamenta en la necesidad de desarrollar y evaluar un enfoque metodológico integral que articule modelos predictivos y criterios sistemáticos de análisis para la detección temprana de abandono estudiantil. Con el propósito de disponer de manera holística variables académicas, conductuales, psicoemocionales e institucionales, superando así las limitaciones de los enfoques tradicionales y aprovechando las oportunidades que ofrecen las nuevas fuentes de recolección de datos educativas.

Palabras clave: deserción universitaria; modelo multivariable; aprendizaje automático

### CONTEXTO

Este proyecto se suscribe en el marco de las iniciativas de investigación aplicada e innovación tecnológica radicada en la UAI, orientado a fortalecer la calidad académica y la gestión basada en datos en la educación superior. La propuesta articula los avances en ciencia de datos, aprendizaje automático y analítica educativa con las necesidades institucionales vinculadas a la retención y permanencia estudiantil.

Con el propósito de integrar de manera holística variables académicas, conductuales, psicoemocionales e institucionales, superando así las limitaciones de los enfoques tradicionales y aprovechando las oportunidades que ofrecen las nuevas fuentes de recolección de datos educativas.

En este sentido, se orienta al estudio de las diferentes propuestas con el propósito de construir un marco estructurado que permita integrar dominios de datos contextuales (sociales, académicos, psicológicos e institucionales) del estudiante para establecer un mecanismo replicable de análisis de estos datos y traducir los resultados en lineamientos operativos para la toma de decisiones.

### LINEAS DE INVESTIGACIÓN y DESARROLLO

La investigación adopta un enfoque cuantitativo, de tipo aplicado y descriptivo-predictivo, dado que busca identificar relaciones causales entre variables, al mismo tiempo, desarrollar un modelo con aplicaciones prácticas en contextos académicos universitarios. Este enfoque se sustenta en el análisis estadístico y técnicas de aprendizaje automático, con el fin de construir y evaluar un modelo predictivo capaz de anticipar la deserción en el primer año. El diseño de la metodología es no experimental, transversal y predictivo, basado en el análisis de datos históricos de estudiantes

### Modelo Multivariable



de primer año, mientras que la muestra se seleccionará a partir de registros académicos de una o más cohortes, considerando criterios de disponibilidad y calidad de los datos.

### RESULTADOS OBTENIDOS / ESPERADOS

Los objetivos de la presente investigación esperan fundamentar la necesidad de desarrollar y evaluar un enfoque metodológico integral que articule modelos predictivos que permita identificar patrones de riesgo asociados a la deserción en el nivel universitario durante los primeros años de la carrera. Para alcanzar esta meta, el modelo propuesto integrará dominios multivariados, con el fin no solo de apoyar a la detección temprana de estudiantes con mayor probabilidad de abandonar la educación superior, sino también facilitar la generación de evidencia formal para la formulación de modelos sostenibles de alerta temprana.

En función de este propósito, el estudio plantea, en primer lugar, realizar un mapeo sistemático de la literatura científica sobre los modelos de deserción estudiantil, poniendo especial atención en los aportes provenientes de diferentes dominios y la interrelación entre ellos. A partir de esta revisión, se analizarán y compararán los distintos modelos teóricos y predictivos reportados, con el objetivo de identificar aquellos que demuestren mayor solidez y respaldo empírico.

Con base en lo anterior, se diseñará y construirá un dataset unificado que integre las variables que han sido consideradas relevantes en investigaciones previas, siguiendo criterios teóricos y metodológicos respaldados por la literatura. Posteriormente, se seleccionarán e implementarán algoritmos de aprendizaje automático acordes con el problema de estudio, y su desempeño se evaluará mediante métricas estandarizadas y pruebas empíricas controladas.

Asimismo, se examinará la capacidad predictiva del modelo desarrollado y su utilidad práctica como herramienta de apoyo para la toma de decisiones académicas. De manera complementaria, se explorarán técnicas de visualización de datos que permitan interpretar y comunicar los resultados de forma clara dentro de contextos institucionales.

Finalmente, a partir de los hallazgos obtenidos, se elaborará un conjunto estructurado de recomendaciones orientadas al diseño e implementación de sistemas de alerta temprana, enfocados en la identificación oportuna de estudiantes con riesgo de deserción.