

AUTORAS

Catalina Cordero Zurita
Sofía Espinoza Gajardo
Arantxa Lezaeta Mac-Lean

PREDICCIÓN DE LA DESERCIÓN ESTUDIANTIL EN CHILE UTILIZANDO DATOS SOCIOECONÓMICOS DE LA ENCUESTA CASEN 2022

Este proyecto tiene como objetivo abordar el problema multifactorial de la deserción escolar en jóvenes chilenos mediante el desarrollo de un modelo predictivo utilizando técnicas de Machine Learning. Usamos el conocido modelo de regresión logística y variables socioeconómicas relevantes de la encuesta CASEN, con el fin de promover intervenciones tempranas que reduzcan las tasas de deserción.



01. INTRODUCCIÓN

La deserción escolar es un problema multifactorial que impacta negativamente el futuro de los jóvenes, reduciendo sus oportunidades laborales y aumentando el riesgo de pobreza. Este proyecto busca enfrentar esta problemática en Chile mediante el desarrollo de una herramienta basada en ciencia de datos y algoritmos predictivos, capaz de identificar tempranamente a adolescentes en riesgo de deserción. Inspirado en literatura y casos internacionales, nuestro objetivo es contribuir al diseño de políticas e intervenciones integrales que fomenten el desarrollo social y económico del país.

03. METODOLOGÍA



Utilizamos la Encuesta de Caracterización Socioeconómica Nacional (CASEN) 2022, una base de datos que proporciona información clave sobre economía, educación, salud y condiciones de vida en Chile.

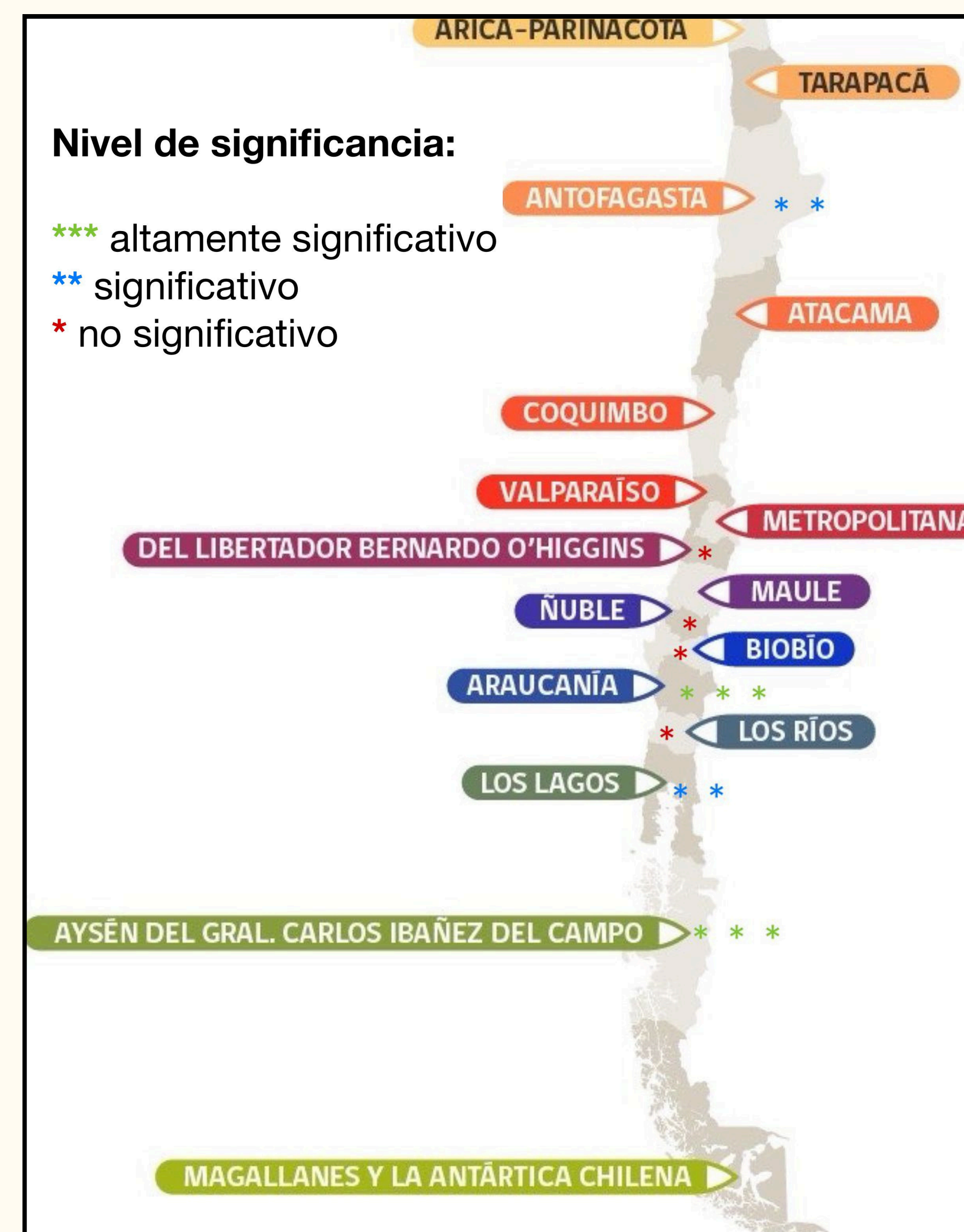
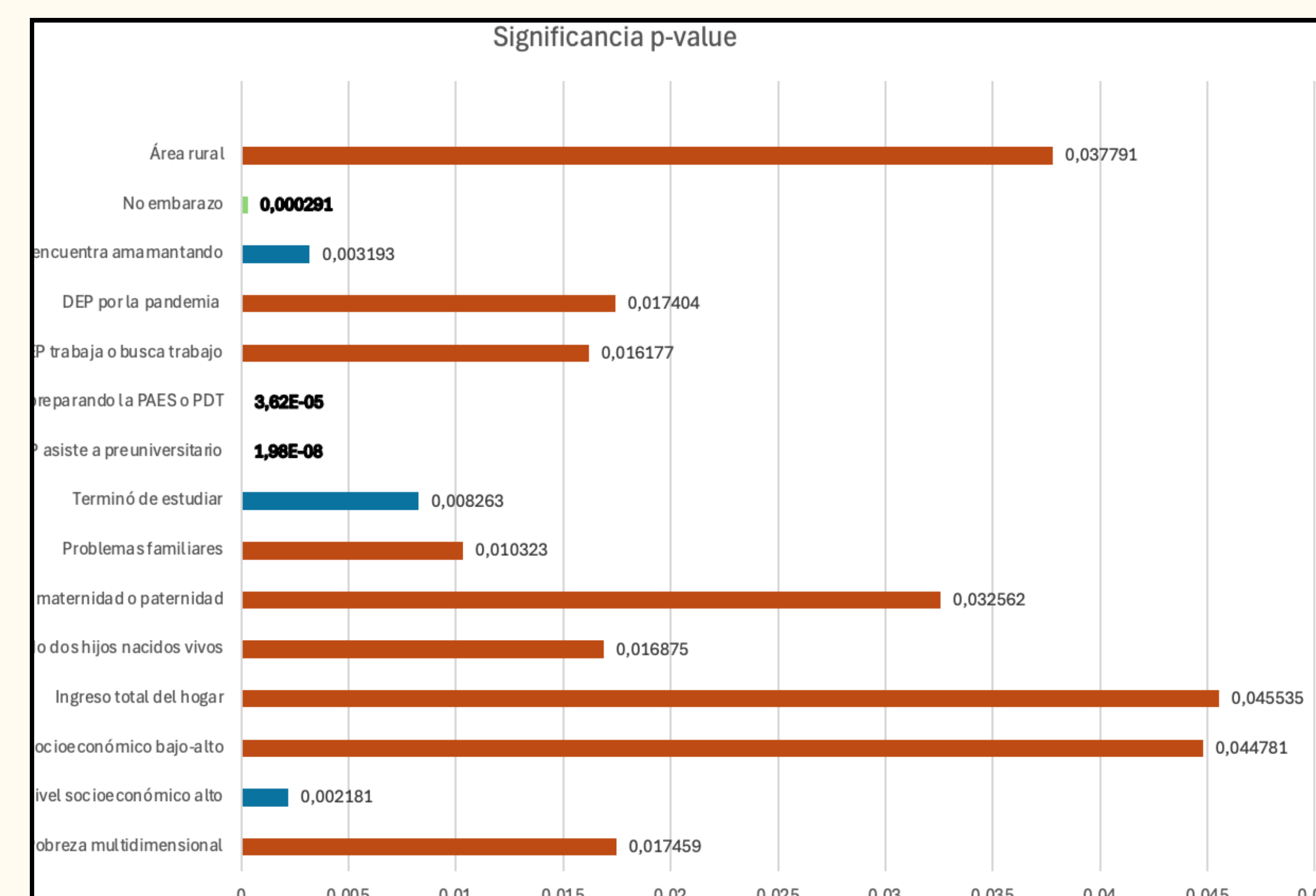
Realizamos una revisión de la literatura para identificar las variables más relevantes, luego analizamos la encuesta y la comparamos con su base de datos para interpretar los resultados.

Algunas de las variables utilizadas en el modelo de regresión logística son:



04. RESULTADOS Y HALLAZGOS

Al aplicar el modelo de regresión logística, se analizaron los niveles de significancia de las variables incluidas, cuyos resultados se presentan gráficamente. Estas visualizaciones destacan las variables más relevantes para la predicción de la deserción escolar, facilitando la interpretación de su impacto en el modelo.



02. OBJETIVOS

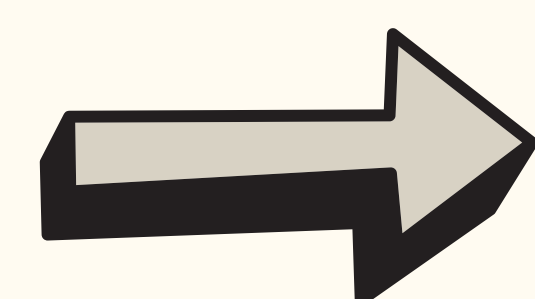
1. Desarrollar un modelo predictivo innovador y preciso para abordar la deserción escolar en Chile.
2. Promover la difusión de conocimiento generado y fortalecer las competencias técnicas del equipo.
3. Escalar el modelo a nivel nacional, buscando influir en políticas públicas y aumentar la retención estudiantil.

05. ANÁLISIS

Variables de impacto:



Realizamos una tabla comparativa con los resultados reales y la predicción del modelo, respecto a la deserción escolar, junto con sus respectivas probabilidades y aciertos de este mismo como se muestra a continuación:



Obtuvimos un test error rate del modelo mediante validación simple del 50.62%. Indicando que el modelo de regresión logística implementado acierta con sus predicciones en solo un $1 - 0,5062 = 49,38\%$ de los casos.

Valor.Real..y_test.	Predicción..y_pred.	probabilidad_desercion	Es_correcto
1	1	0.6622939	TRUE
1	1	0.6879369	TRUE
1	1	0.5258183	TRUE
1	1	0.8328215	TRUE
1	1	0.6862945	TRUE
1	1	0.8366232	TRUE
0	1	0.5674499	FALSE

CONTACTO

¿Tienes alguna pregunta? ¿Te gustaría contactarnos o saber más sobre el proyecto? Para esto y más, visita nuestra página web escaneando el siguiente código QR:

