

Modelo predictivo sobre las estadísticas de educación media en Colombia

Juan Julio y Camilo Sinning

Junio 8, 2022

Introducción

Esta presentación se enfocará en exponer el proyecto de analítica de datos.

Base de datos

https://www.datos.gov.co/Educaci-n/MEN_ESTADISTICAS_EN_EDUCACION_EN_PREESCOLAR-B-SICA/nudc-7mev

El proyecto de analítica de datos tiene la finalidad de aprobar o rechazar hipótesis cuyo eje central es la educación.

Pregunta problema

¿Existe alguna relación entre los porcentajes de aprobación de un municipio, su departamento y la deserción escolar en educación media?

Hipótesis

Los departamentos con mayor índice de ruralidad son aquellos en los cuales se presenta una mayor deserción.

Objetivos

Objetivo general

Desarrollar un modelo predictivo con el cual se pueda pronosticar la deserción en la educación media de cierto municipio.

Objetivos específicos

- Analizar datos provenientes de la base de datos
- Establecer el modelo predictivo
- Demostrar la veracidad de la hipótesis

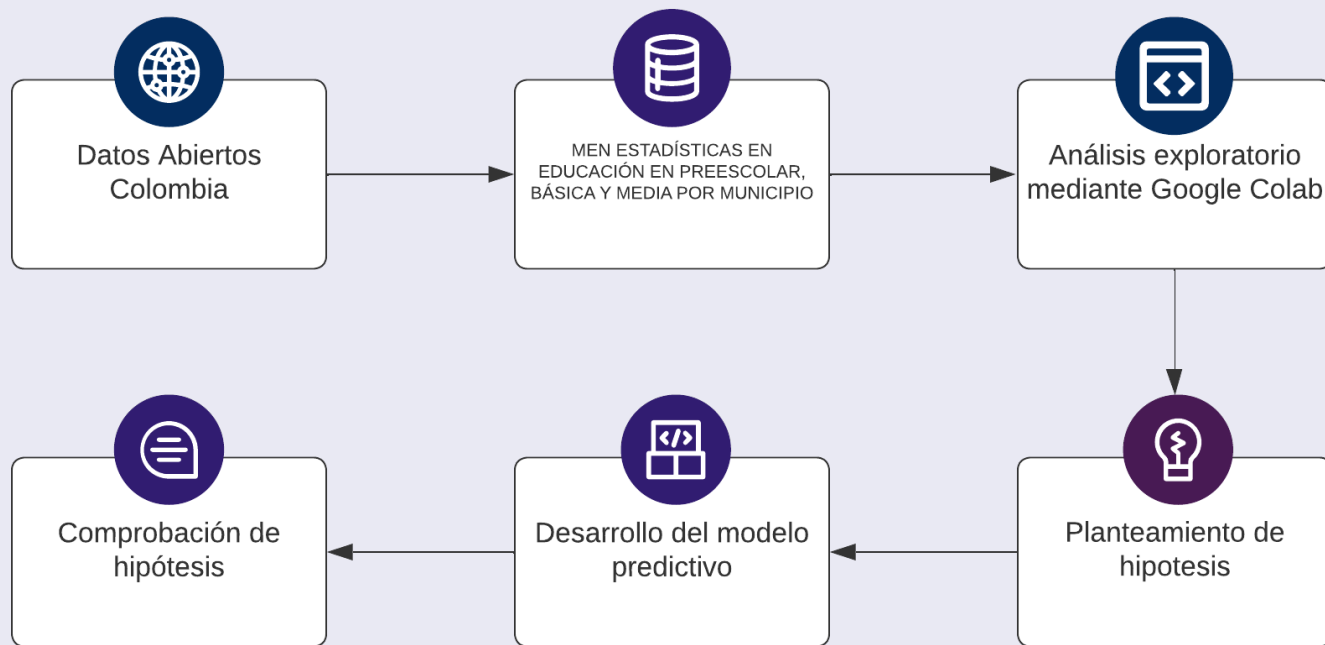


Figure 1: Metodología de la solución

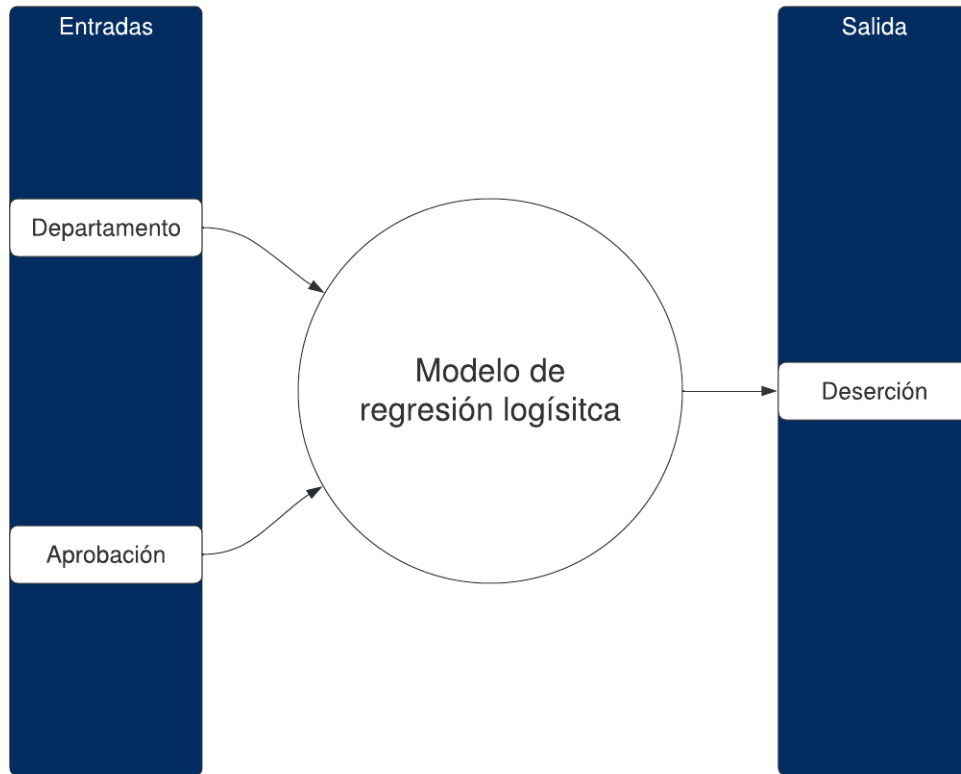


Figure 3: Modelo

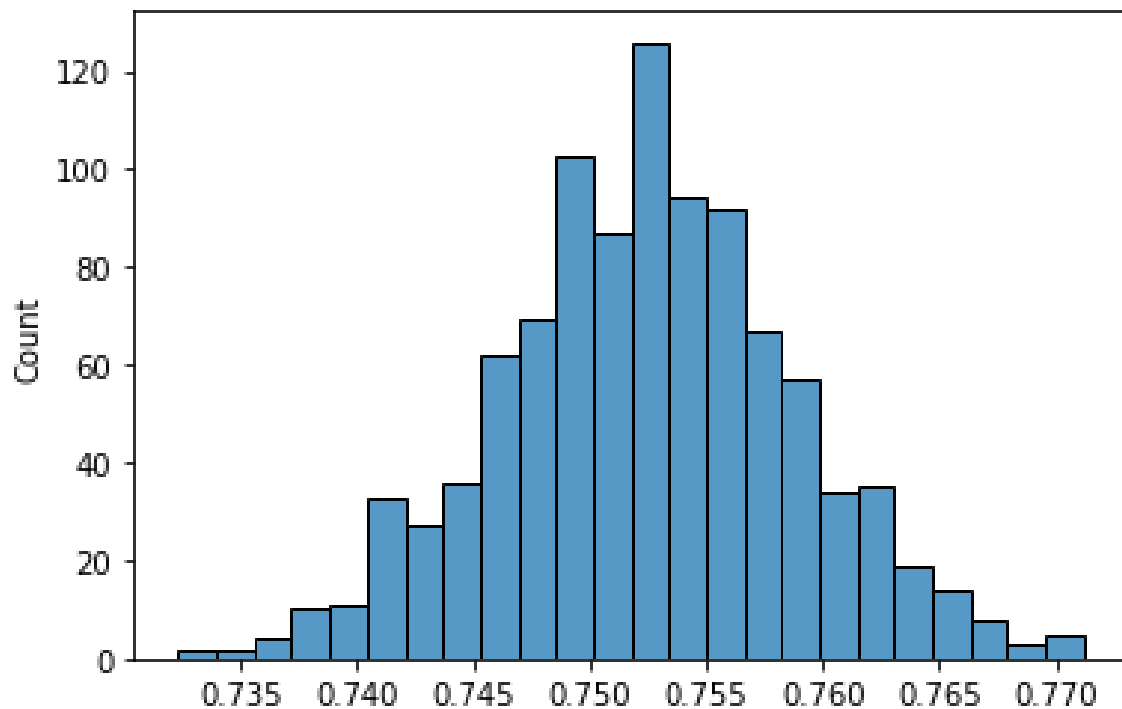


Figure 4: Gráfico de puntaje del modelo

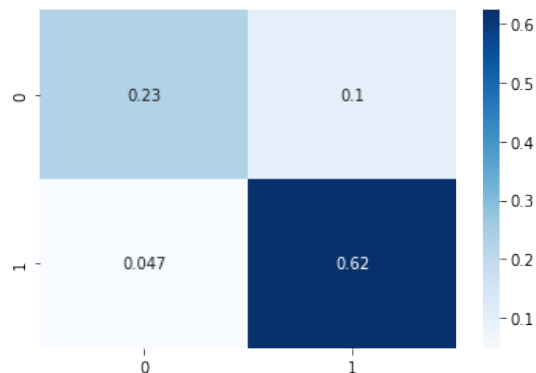
Análisis

Score y varianza

En el desarrollo del modelo predictivo para estimar la deserción se obtuvo un puntaje promedio en 1000 iteraciones de 75% aproximadamente y una varianza de 4%.

Hipótesis

Usando los departamentos de Atlántico, Bogotá D.C., Caldas, Guainía, Islas, Putumayo, Vaupés, Vichada. los datos indican que en un 85% el modelo aprobó la hipótesis como se muestra en el siguiente gráfico:.



El proyecto de analítica de datos permitió:

- ① Usar el modelo de predicción para responder la pregunta problema
- ② Apoyándose en los datos de ruralidad provistos por Naciones Unidas comparados con el índice de deserción, para aprobar la hipótesis [1]



C. Bogota, J. C. Ramírez, and J. M. de Aguas. Naciones Unidas, 2017, vol. 1.