

PROYECTO

PRIMERA ENTREGA

Presentado por:
KAREN ELIANA GARCIA PÉREZ

Curso:
Introducción a la inteligencia artificial

Docente:
Raul Ramos Pollan



UNIVERSIDAD DE ANTIOQUIA
FACULTAD DE INGENIERÍA
MEDELLÍN 2021

Planteamiento del problema

Predicción de deserción de estudiantes, se busca comprender e identificar los factores de riesgo de deserción de los estudiantes de una institución de educación superior, para así desde el área correspondiente de la institución de educación superior, realizar intervenciones tempranas para mejorar la tasa de permanencia de los mismos

Dataset que se va a usar

Se hará uso del data set de kaggle <https://www.kaggle.com/datasets/thedevastator/higher-education-predictors-of-student-retention?resource=download> el cual contiene el registro de 4424 estudiantes y 35 columnas, de las cuales 18 son categóricas, entre las cuales encontramos, datos demográficos, factores socioeconómicos e información sobre el rendimiento académico en columnas tales como:

- Estado civil
- Cualificación previa
- Nacionalidad
- Necesidades educativas especiales
- Deudor
- Edad de inscripción
- Tasa de desempleo
- Objetivo
- Profesión de la madre
- Unidades curriculares 1er sem (matriculados)
- Unidades curriculares 1er sem (aprobado)

Métricas

La métrica de evaluación principal para el modelo será la Exactitud dada por : **Exactitud** = $(TP+TN)/total$, donde:

- **TP** : Verdadero Positivo (Predicho Verdadero y Verdadero en realidad)
- **TN** : Verdadero Negativo (Predicho Falso y Falso en realidad)

En cuanto a las métricas para el negocio se espera que el porcentaje de acierto sea superior al 70%

Criterio sobre el cual sería el desempeño deseable en producción.

En este caso se espera obtener con alto porcentaje de acierto las condiciones y factores que aumentan o disminuyen la deserción estudiantil en una institución de educación superior