

Carrera: Tecnicatura Superior en Ciencia de Datos e Inteligencia Artificial.

Institución: Politécnico Malvinas Argentinas

Materia: Aprendizaje Automático

Proyecto Aprendizaje Automático: Rendimiento Académico de los Estudiantes

Alumna: Barrios, Yanina Soledad

Profesor de la Catedra: Mirabete, Martin

Entrega 2: Descripción del Dataset y Origen

- En esta segunda entrega, debe tener acceso al dataset que utilizará en su proyecto de [Aprendizaje Automático](#).
- Proporcione una descripción completa del dataset, incluyendo la cantidad de instancias, características (columnas), tipos de datos, y cualquier información relevante.
- Informe sobre el origen del dataset, es decir, de dónde provienen los datos. Esto puede incluir la fuente, la fecha de adquisición y cualquier proceso de recopilación o preprocesamiento que haya realizado.
- Se recomienda utilizar dataset de dominio público (ipiec, datos.gob.ar, etc).
- Esta entrega debe ser realizada en el repositorio GIT, incluyendo archivos de dataset y documentos.

Descripción general:

Este dataset contiene información sobre el desempeño académico de estudiantes, junto con variables relacionadas con su entorno social, familiar y educativo. Su propósito es servir como base para predecir el rendimiento de los estudiantes utilizando técnicas de aprendizaje automático, como clasificación o regresión.

Descripción del Dataset y Origen

Nombre del Dataset: Rendimiento Académico de los Estudiantes

Origen:

- **Fuente:** Kaggle – [Student Academic Performance Dataset](#)
- **Autor:** Rohith Anand
- **Fecha de adquisición:** Mayo 2025
- **Licencia:** Dominio público / uso educativo (según política de Kaggle)

Estructura del Dataset:

El archivo Dataset.xlsx contiene una única hoja con datos estructurados

- * Cantidad de instancias (filas): 1000
- * Cantidad de características (columnas): 15

Listas de columnas y tipos de datos:

student_id : Identificador único del estudiante (Categórico)

name: Nombre del estudiante (Categórico)

gender: Género del estudiante (Male/ Female) (Categórico)

age: Edad del estudiante (Numérico(entero))

grade_level: Nivel educativo o curso (ej., 9th, 10th) (Categórico)

math_score: Calificación en Matemática (Numérico(entero))

reading_score: Calificación en Lectura (Numérico(entero))

writing_score: Calificación en Escritura (Numérico(entero))

attendance_rate: Porcentaje de asistencia del estudiante (Numérico(decimal))

parent_education: Nivel educativo de los padres (ej., High School, Graduate)

study_hours: Cantidad de hora de estudio por día (Categórico)

internet_access: Acceso a internet en casa (Yes/ No) (Categórico)

lunch_time: Tipo de almuerzo recibido (Standard/ Free) (Categórico)

extra_activities: Participa en actividades extracurriculares (Yes/No) (Categórico)

final_result: Resultado final del estudiante (Pass/ Full) (Categórico)

Preprocesamiento: Hasta el momento, no se ha realizado preprocesamiento específico sobre los datos.

Dirección del Github: <https://github.com/Yansol23/Alumnos>