**Ordenar los números de las slides**

Script

*¿Una cosa es la feature que más afecte sobre la nota y otra es la que mejor la prediga🡪 es lo mismo?*

¿Para qué predecir un nota de exámenes? Sirve para los dos lados. Saber qé es lo que más influye y también sabiendo esto poder predecir.

### Cuál es la situación actual del aprendizaje en Alemania

If we found the most predictive features of low math result…

### El profe puede ayudar a los alumnos que tienen estas características

Prediciendo cuales son los alumnos que tienen más probabilidad de sacar mala nota en matemáticas.

### A qué grupos es necesario apoyar para levantar la nota de matemática

Por ejemplo: Chicas. De una cierta etnia, con padres que…, hacen deporte..

Para que Alemania pueda subir en el ranking de matemáticas necesitamos dar apoyo a\_\_\_

Elección del modelo🡪¿Cómo hago para encontrar el mejor parámetro? 🡪Elijo primero el mejor modelo.

Modelo🡪¿Cúal es el mejor modelo para predecir la mejor feature?

Resultados🡪¿Cuál es el mejor parámetro/característica?

# Portada

## What do I need to do to get an A+ in math?

### What is the most important feature in math scores?

#### Is it possible to predict student grades based on these variables?

**Topic introduction**

## Why am I doing this?

Aproaches:

· La matemática se ha hecho con el control de todo, Introducción del tema

· Hijo de un mago

· A los padres solo les interesa la nota de matemáticas

· Ya no hace falta saber leer, solo matemáticas

· Competición entre matemáticas y lengua

· Resultados de matemáticas en la EU.

***Mathematics results in the World/Germany***

Agtung!

Alemania está perdiendo el liderazgo en matemática en el mundo. Necesitamos encontrar la forma de dar vuelta esto.

### For this we need to discover which is the feature that predict (impacts) the most.

Dónde debería invertir el estado alemán su dinero para mejorar este indicador.

## Aim

For this reason I set out to find the characteristic that most influences the mathematics grade.

## Índice/orden de la presentación

Introducción

EDA

Regressiones

## Dataset

(una o dos slides)

Para eso he buscado un dataset en Kaggle y he encontrado el siguiente.

El dataset goza de 19243 rows después de limpiarlo.

Tiene las siguientes categorías

**Gender**: Gender of the student (male/female)

**EthnicGroup**: Ethnic group of the student (group A to E)

**ParentEduc**: Parent(s) education background (from some\_highschool to master's degree)

**LunchType**: School lunch type (standard or free/reduced)

**TestPrep**: Test preparation course followed (completed or none)

**ParentMaritalStatus**: Parent(s) marital status (married/single/widowed/divorced)

**PracticeSport**: How often the student parctice sport (never/sometimes/regularly))

**IsFirstChild**: If the child is first child in the family or not (yes/no)

**NrSiblings**: Number of siblings the student has (0 to 7)

**TransportMeans**: Means of transport to school (schoolbus/private)

**WklyStudyHours**: Weekly self-study hours(less that 5hrs; between 5 and 10hrs; more than 10hrs)

**MathScore**: math test score(0-100)

**ReadingScore**: reading test score(0-100)

**WritingScore**: writing test score(0-100)

**Gender**, **EthnicGroup, ParentEduc, LunchType, TestPrep, ParentMaritalStatus, PracticeSport, IsFirstChild, NrSiblings, TransportMeans, WklyStudyHours, MathScore, ReadingScore, WritingScore**

## Analysis

## E.D.A.

Los gráficos

## Regresiones

Linear regression

Decision tree

KNN

Una slide para cada una con su resultado

Una con el resumen de las tres en la que se compare el R2 de los tres.

Conclusión: Respondiendo a la pregunta del principio.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Linear  Regression | Decision  Tree | KNN |
| R2 | 0.842 | 0.83 - 0.82 | 0.80 - 0.77 |