

INFORME DEL MÉTODO DE LA INGENIERÍA

IDENTIFICACIÓN DEL PROBLEMA

Identificación de necesidades

- Mostrar una tabla en la cual se pueda apreciar la información de cada uno de los estudiantes. Además, tener la posibilidad de filtrar esta tabla mediante los campos de la base de datos.
- Mostrar los 5 factores más representativos de la información en un gráfico cada uno.
- Predecir las notas de los exámenes finales de futuros estudiantes, clasificándolos en base a la información que se tiene del dataset.

Definición del problema

El departamento de educación de Lisboa, Portugal ha decidido realizar una investigación frente al desempeño académico de los estudiantes de secundaria de la ciudad. Para esto ha tomado una muestra de dos escuelas: Gabriel Pereira y Mousinho da Silveira.

El objetivo del departamento con este estudio es comprender cómo afectan los distintos factores tales como: Tipo de asentamiento (Urbano o rural), nivel de educación de los padres, tiempo semanal dedicado a estudiar, etc. Con esta información se busca poder predecir la nota que posteriores estudiantes teniendo situaciones similares podrían obtener en su examen final.

Por último, han manifestado que les gustaría ver gráficamente la información que han recolectado utilizando tablas y diagramas.

RECOPILACIÓN DE INFORMACIÓN

Definiciones

Clasificación:

Relación de los clasificados en una determinada prueba.¹

Clasificar:

Ordenar o disponer por clases algo.²

Clasificación (ML): Proceso de categorizar un set de datos en clases, puede ser realizado en datos estructurados o no estructurados. El proceso empieza prediciendo la clase de puntos de datos. Las clases son comúnmente referidas como objetivo, marcador o categoría.³

Base de datos:

Una base de datos es una colección organizada de información estructurada, o datos, típicamente almacenados electrónicamente en un sistema de computadora.⁴

Diagrama de barras:

Es un gráfico que se utiliza para representar datos de variables cualitativas o discretas. Está formado por **barras** rectangulares cuya altura es proporcional a la frecuencia de cada uno de los valores de la variable.⁵

¹ "clasificación | Definición | Diccionario de la lengua española | RAE"
<https://dle.rae.es/clasificaci%C3%B3n>.

² "clasificar | Definición | Diccionario de la lengua española | RAE" <https://dle.rae.es/clasificar>.

³ "Classification In Machine Learning | Classification ... - Edureka." 21 jul. 2020,
<https://www.edureka.co/blog/classification-in-machine-learning/>.

⁴ "¿Qué es una base de datos? | Oracle Colombia."
<https://www.oracle.com/co/database/what-is-database/>.

⁵ "Diagrama de barras - Universo Formulas." 14 abr. 2014,
<https://www.universoformulas.com/estadistica/descriptiva/diagrama-barras/>. Se consultó el 9 abr. 2021.

Diagrama de pastel:

Un diagrama de pastel es un círculo dividido en partes, donde el área de cada parte es proporcional al número de datos de cada categoría.⁶

Diagrama de puntos:

Un diagrama de puntos es una gráfica utilizada para ilustrar un número reducido de datos, la cual permite identificar con facilidad dos características:

1. La localización de los datos.
2. La dispersión o variabilidad de los datos.

Este diagrama muestra cada uno de los elementos de un conjunto de datos numéricos por encima de una recta numérica.⁷

⁶ "GRÁFICA DE PASTEL - UNAM."

<http://asesorias.cuautitlan2.unam.mx/Laboratoriovirtualdeestadistica/DOCUMENTOS/TEMA%201/5.%20GRAFICA%20DE%20%20PASTEL.pdf>. Se consultó el 9 abr. 2021.

⁷ "Diagrama de Puntos - Temas de Estadística." 16 ene. 2013, <http://tsu-estadistica.blogspot.com/2013/01/diagrama-de-puntos.html>. Se consultó el 9 abr. 2021.