PROYECTO FINAL DE INTELIGENCIA ARTIFICIAL- NIVEL INTEGRADOR

MODELO PREDICTIVO DE LA RESISTENCIA DEL CEMENTO

Julián David Pérez Navarro

Juan Sebastián Jaramillo

Marcela Aristizábal Ríos

Juan Carlos Torres

Kevin Julián Chavarria Olarte

TALENTO TECH

UNIVERSIDAD DE ANTIOQUIA- UNIVERSIDAD DE CALDAS

2024

**Table of Contents**

[**1.**](#_heading=h.gjdgxs) **PLANIFICACIÓN DEL PROYECTO 3**

[**1.1**](#_heading=h.30j0zll) **REUNIÓN INICIAL. 3**

[**1.2**](#_heading=h.2et92p0) **DEFINICIÓN DE OBJETIVOS 4**

[**1.3**](#_heading=h.tyjcwt) **IDENTIFICACIÓN DE DATOS NECESARIOS 4**

[**2.**](#_heading=h.3dy6vkm) **PREPROCESAMIENTO DE DATOS 5**

[**2.1**](#_heading=h.1t3h5sf) **LIMPIEZA DE DATOS 5**

[**2.2**](#_heading=h.4d34og8) **NORMALIZACIÓN Y SELECCIÓN DE CARACTERÍSTICAS 5**

[**3.**](#_heading=h.2s8eyo1) **DESARROLLO DE MODELOS DE IA 6**

[**3.1**](#_heading=h.17dp8vu) **DEFINICIÓN DE LA ARQUITECTURA DEL MODELO 6**

[**3.2**](#_heading=h.3rdcrjn) **IMPLEMENTACIÓN DEL MODELO 6**

[**3.3**](#_heading=h.26in1rg) **PRIMERA EVALUACIÓN 6**

[**4.**](#_heading=h.lnxbz9) **EVALUACIÓN DE RENDIMIENTO 7**

[**4.1**](#_heading=h.35nkun2) **AJUSTE DE HIPERPARÁMETROS 7**

[**4.2**](#_heading=h.1ksv4uv) **APLICACIÓN DE REGULARIZACIÓN 7**

[**4.3**](#_heading=h.44sinio) **IMPLEMENTACIÓN DE EARLY STOPPING 7**

[**5.**](#_heading=h.2jxsxqh) **EVALUACIÓN FINAL Y RESULTADOS 8**

[**5.1**](#_heading=h.z337ya) **EVALUACIÓN DEL MODELO 8**

[**5.2**](#_heading=h.3j2qqm3) **VISUALIZACIÓN DE LOS RESULTADOS 8**

[**5.3**](#_heading=h.1y810tw) **DOCUMENTACIÓN 8**

[**6.**](#_heading=h.4i7ojhp) **PRESENTACIÓN DEL PROYECTO 9**

[**6.1**](#_heading=h.2xcytpi) **PREPARACIÓN DE LA PRESENTACIÓN 9**

[**6.2**](#_heading=h.1ci93xb) **PRESENTACIÓN AL GRUPO 9**

[**6.3**](#_heading=h.3whwml4) **DISCUSIÓN Y FEEDBACK 9**

1. PLANIFICACIÓN DEL PROYECTO

MARCO CONCEPTUAL

**Cemento**: Es un material en polvo que actúa como aglutinante. Al mezclarse con agua, reacciona químicamente (hidratación) y se endurece. Es uno de los componentes principales en la elaboración del concreto, pero por sí solo no se utiliza como material estructural.

**Concreto**: Es un material compuesto formado al mezclar cemento, agua, arena (agregado fino) y grava (agregado grueso). El cemento une los agregados, formando una masa rígida y durable cuando se seca. El concreto es el material final usado en construcciones, como edificios y puentes.

Diferencia clave: El cemento es un componente del concreto, mientras que el concreto es el producto final usado en la construcción.

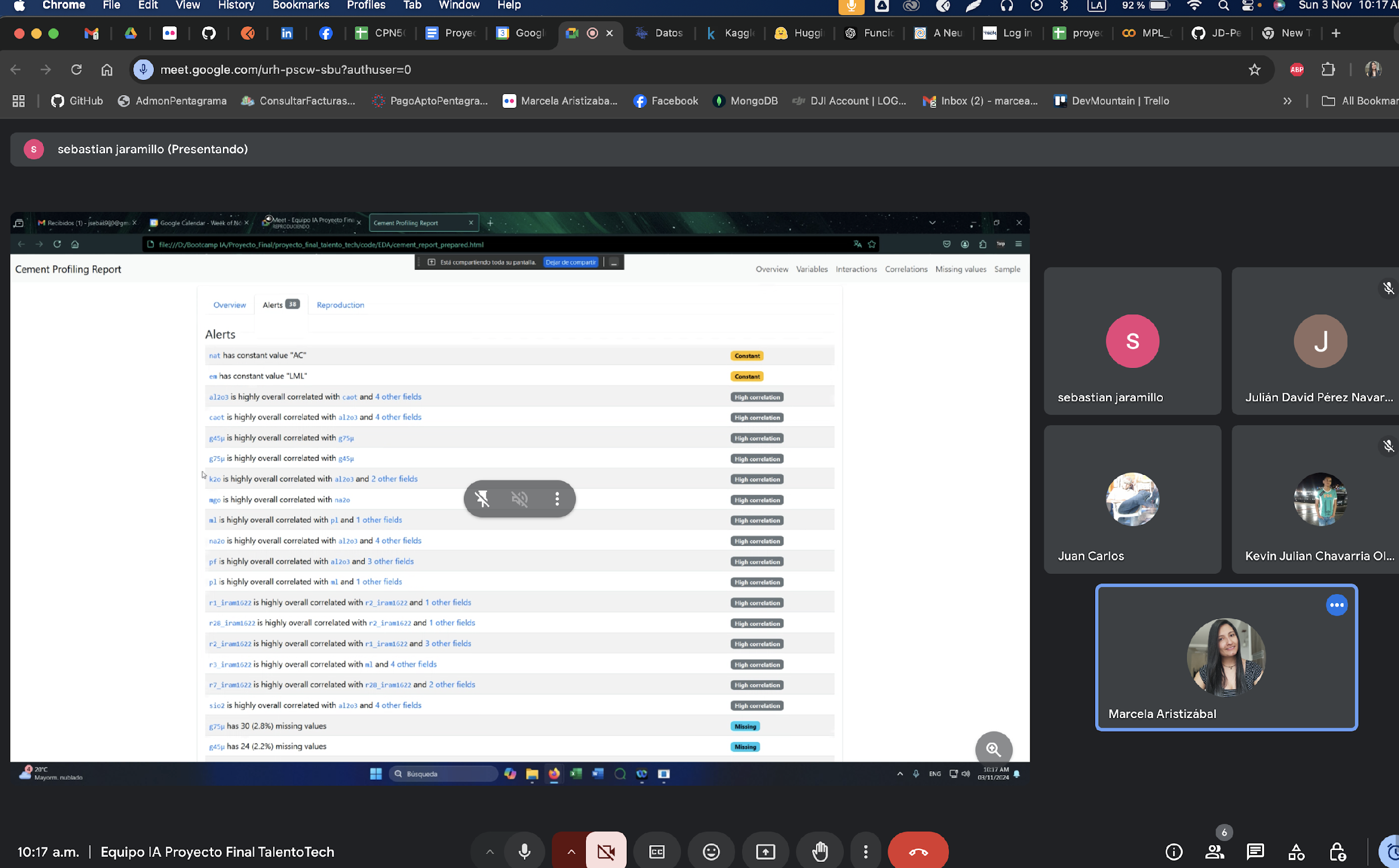
**Resistencia del cemento:** La resistencia del cemento se refiere a su capacidad para soportar fuerzas o cargas sin romperse o deformarse permanentemente. Esta resistencia es fundamental en la construcción, ya que determina la estabilidad y durabilidad de las estructuras hechas con concreto o mortero que utilizan cemento como aglutinante.

ALCANCE

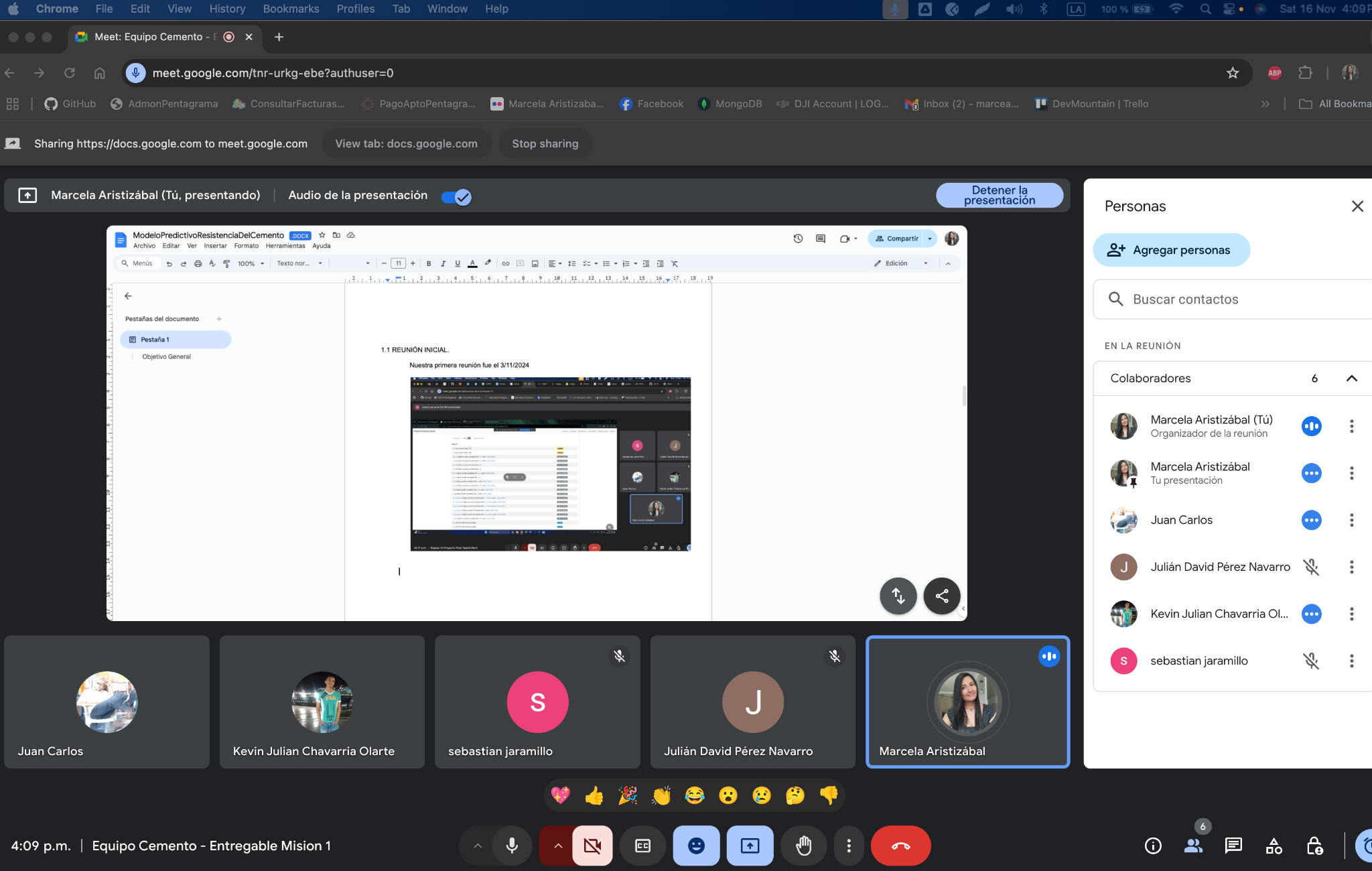
El proyecto se centrará en el desarrollo de un modelo predictivo para estimar la resistencia del cemento en función de su composición química, utilizando técnicas de machine learning y deep learning que estimen su resistencia en días clave.

* 1. REUNIÓN INICIAL.

Nuestra primera reunión fue el 3/11/2024



Segunda reunión 16/11/2024 ajustes: definición de alcance, objetivos generales y específicos



* 1. DEFINICIÓN DE OBJETIVOS

### **Objetivo General**

Desarrollar un modelo predictivo utilizando machine learning para estimar la resistencia del cemento en días clave, basándose en la composición química de la mezcla.

**Objetivos Específicos**

**1 Objetivo:** Recolectar y preparar los datos:

**Acciones:** Obtener un conjunto de datos amplio que contenga las características químicas del cemento y las mediciones de resistencia a lo largo del tiempo.

**2 Objetivo:** Seleccionar y crear el modelo

**Acciones**: Implementar tanto modelos de machine learning clásicos como redes neuronales para predecir la resistencia del cemento en los días clave.

**3 Objetivo:** Entrenar y validar el modelo:

**Acciones**: Entrenar los modelos utilizando los datos disponibles y validar su capacidad predictiva.

**4 Objetivo:** Generar la ecuación predictiva:

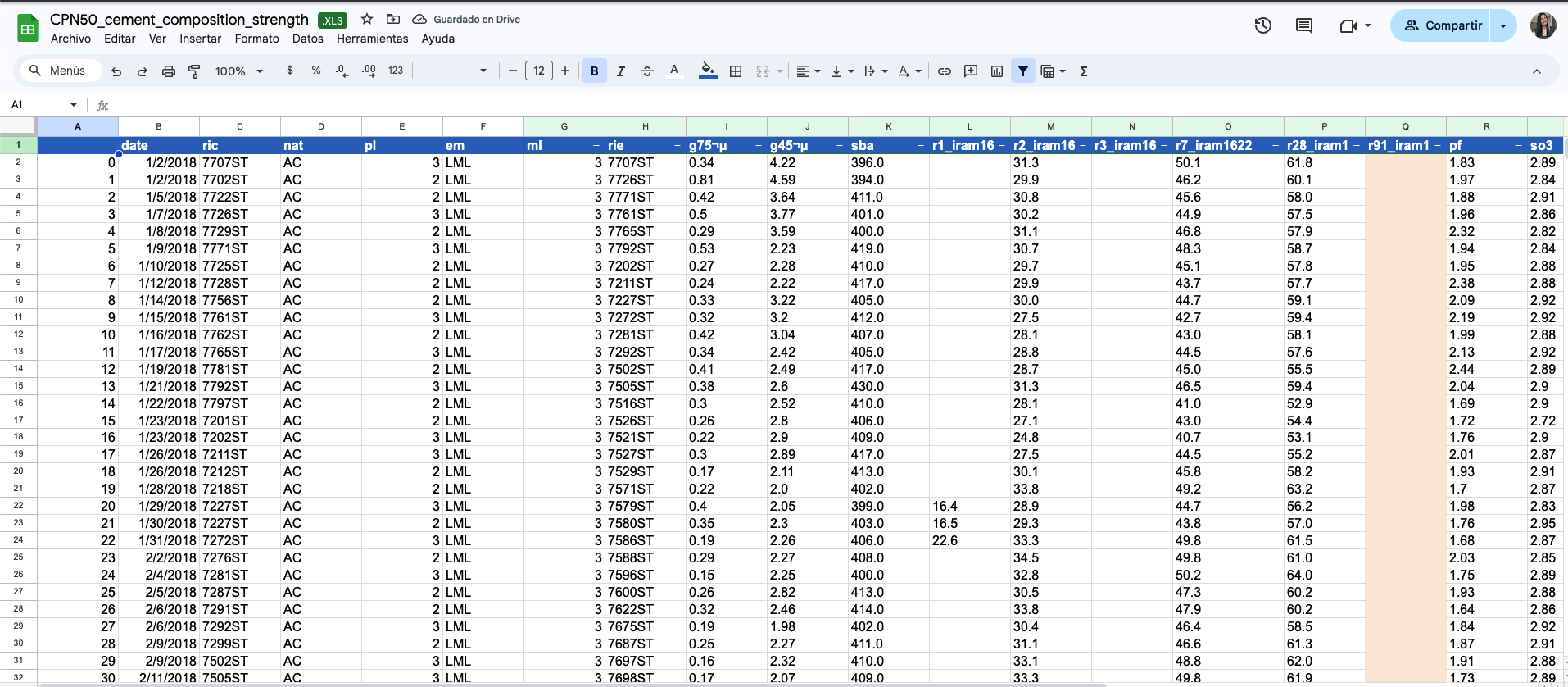
**Acciones**: Derivar una ecuación matemática que relacione las características químicas del cemento con la resistencia en los días clave.

**5 Objetivo:** Evaluar el rendimiento del modelo:

**Acciones**: Evaluar el modelo predictivo de manera exhaustiva y proponer mejoras si es necesario.

* 1. IDENTIFICACIÓN DE DATOS NECESARIOS

Dataset “CPN50\_cement\_composition\_strength.xls”



1. PREPROCESAMIENTO DE DATOS
   1. LIMPIEZA DE DATOS

Se utilizó la librería **KNNImputer** que permite imputar (rellenar) los valores faltantes de una manera más robusta y precisa en comparación con métodos simples como la media o mediana, ya que utiliza la similitud entre observaciones.

Se utilizó **KNNImputer**, con el fin de preservar las relaciones y patrones inherentes en los datos, resultando en un conjunto de datos más representativo para entrenar el modelo.

Link del Repositorio en el que estamos trabajando:

[**https://github.com/JD-PerezN/proyecto\_final\_talento\_tech**](https://github.com/JD-PerezN/proyecto_final_talento_tech)

* 1. NORMALIZACIÓN Y SELECCIÓN DE CARACTERÍSTICAS

Del dataset “CPN50\_cement\_composition\_strength.csv” se realizó lo siguiente:

* La columna ``r91\_iram1622`` no tiene datos, por lo que se eliminará del conjunto de datos
* Las columnas ``g75µ`` y ``g45µ`` están altamente correlacionadas, y viendo los valores faltantes y ceros, se eliminará la columna ``g75µ`` y nos quedaremos con ``g45µ`` para una posterior imputación de los datos
* Por definiciones de negocio, las columnas ``date``, ``ric``, ``nat``, ``pl``, ``em``, ``ml`` y ``rie`` no aportan valor para el modelado de los datos, ya que corresponden a fechas, molinos, identificadores
* Las columnas características que tengan valores faltantes se imputarán usando KNNImputer debido a su robustez
* Se evaluará la posibilidad de modelar las resistencias usando modelos de machine learning

1. DESARROLLO DE MODELOS DE IA
   1. DEFINICIÓN DE LA ARQUITECTURA DEL MODELO
   2. IMPLEMENTACIÓN DEL MODELO
   3. PRIMERA EVALUACIÓN
2. EVALUACIÓN DE RENDIMIENTO
   1. AJUSTE DE HIPERPARÁMETROS
   2. APLICACIÓN DE REGULARIZACIÓN
   3. IMPLEMENTACIÓN DE EARLY STOPPING
3. EVALUACIÓN FINAL Y RESULTADOS
   1. EVALUACIÓN DEL MODELO
   2. VISUALIZACIÓN DE LOS RESULTADOS
   3. DOCUMENTACIÓN
4. PRESENTACIÓN DEL PROYECTO
   1. PREPARACIÓN DE LA PRESENTACIÓN
   2. PRESENTACIÓN AL GRUPO
   3. DISCUSIÓN Y FEEDBACK