



# Relicia

## Clasificación numismática

Vanessa López Aguilera  
Febrero 2025

# Introducción

- Problema.
- Solución.
- Origen de los datos.
- Modelos de Machine Learning:
  - Obtención de los datos.
  - Preprocesamiento de datos.
  - Modelos:
    - Supervisados.
    - No supervisados.
- Escalabilidad.

# Problema

Estado de las monedas cuando se encuentran en las excavaciones.



Yacimiento de Marea / Universidad de Varsovia



Monedas halladas en la Cueva de las Cabras / Universidad Santiago de Compostela

# Solución

Clasificar monedas antiguas por su época histórica usando  
Machine Learning.

TARGET:

- Épocas históricas:
  - Edad del Hierro (s.IX a.C. - s.I d.C.)
  - Época romana (s.I - s.V d.C.).
  - Época bajomedieval (s.V - s.XI d.C.)
  - Época medieval (s. XI - XVI d.C.)
  - Época postmedieval (s. XVI - s.XVII d.C.)

# Solución

Clasificar monedas antiguas por su época histórica usando  
Machine Learning.

## FEATURES:

- Espesor.
- Diámetro.
- Peso.
- Axis.
- Axis\_conocido.
- Materia prima.
- Método de fabricación.
- Imágenes.



# Origen de los datos

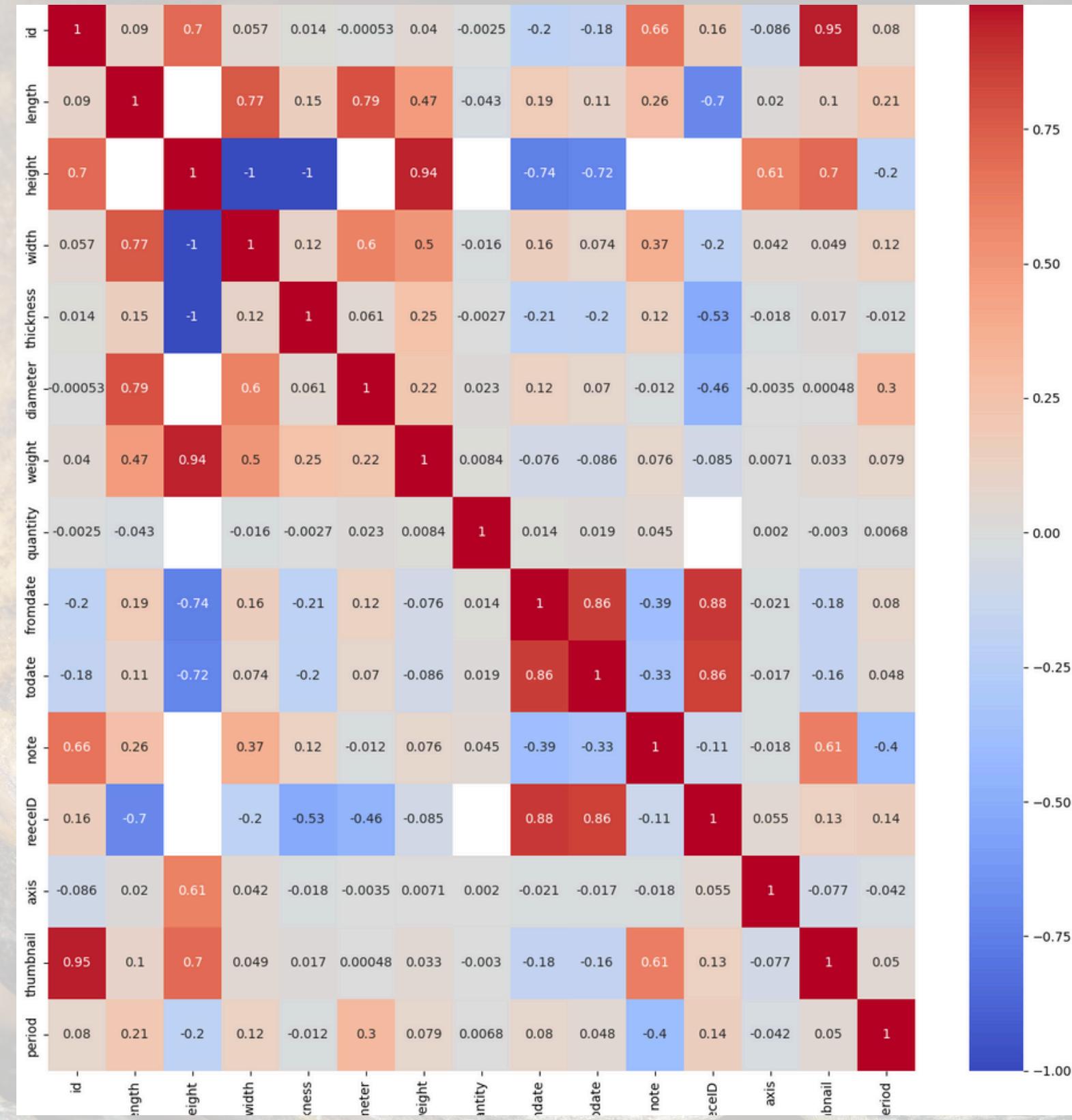


## 1. Datasets procedente de Kaggle:

- 5 datasets provenientes de 'Coin Images from the Portable Antiquities Scheme' subido por Sarah Good (<https://www.kaggle.com/segood>).



# Origen de los datos





# Origen de los datos



## 2. Webscrapping:

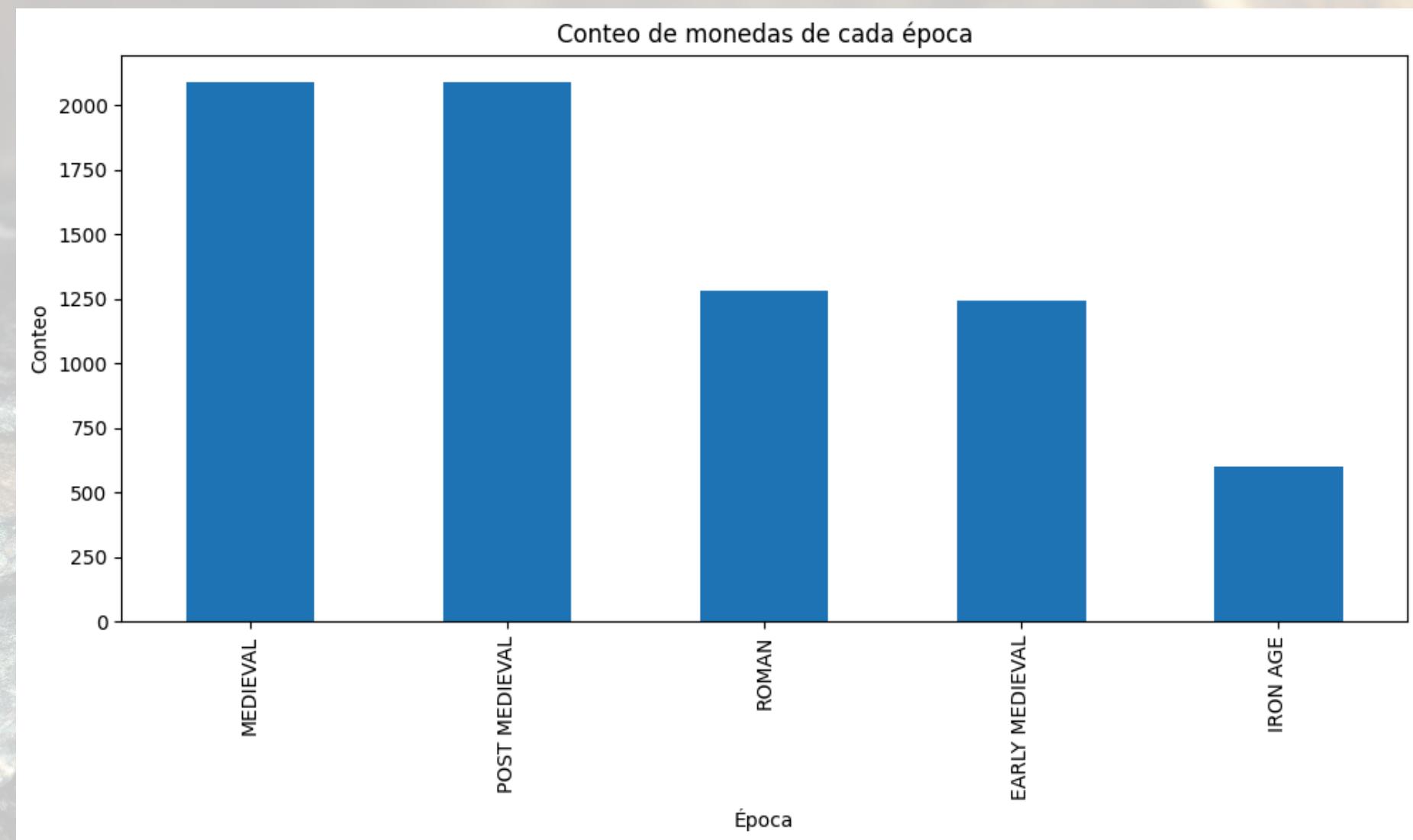
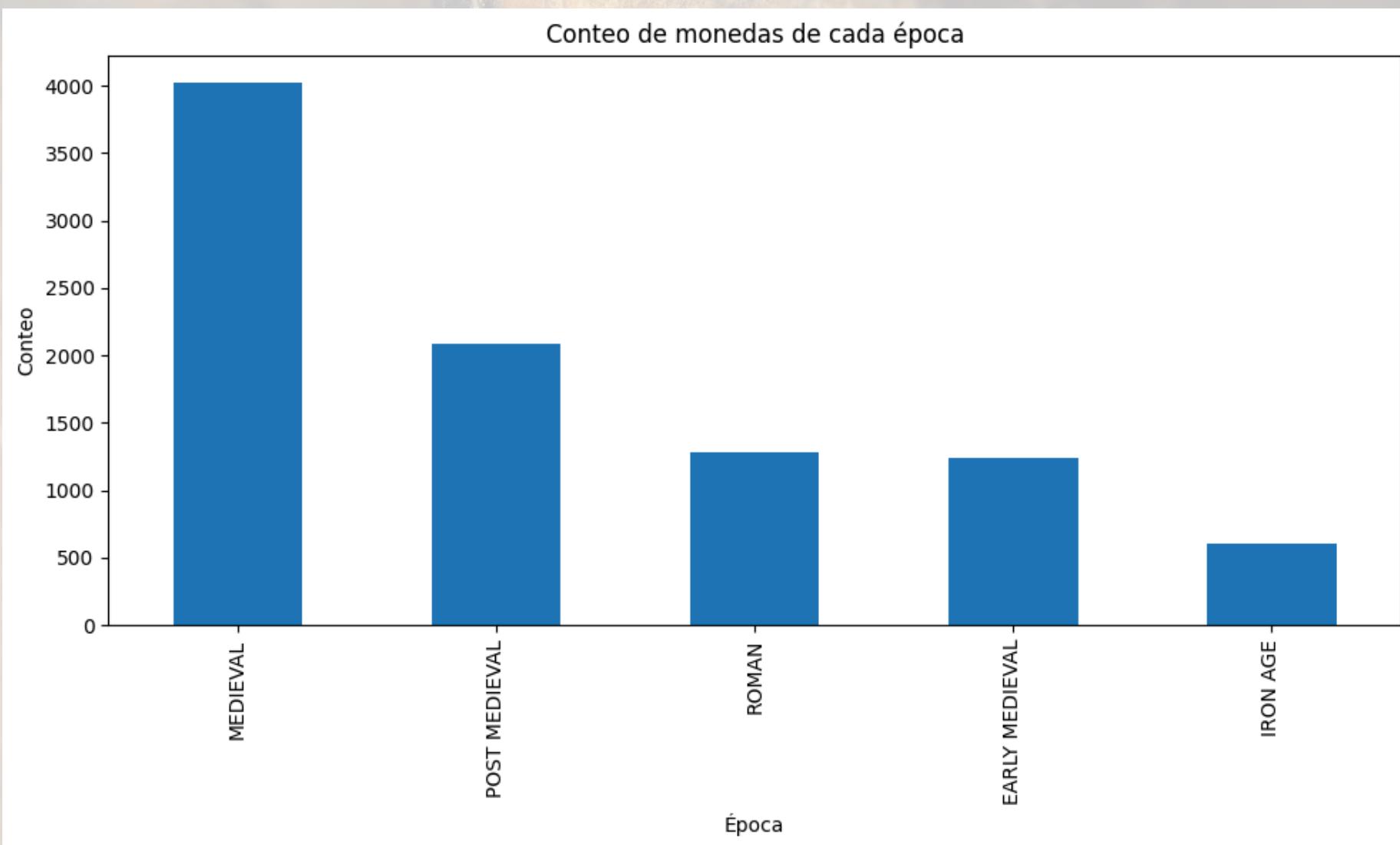
- <https://finds.org.uk/database> perteneciente al The British Museum.





# Procesamiento de los datos

## 1. Balancear el dataset.

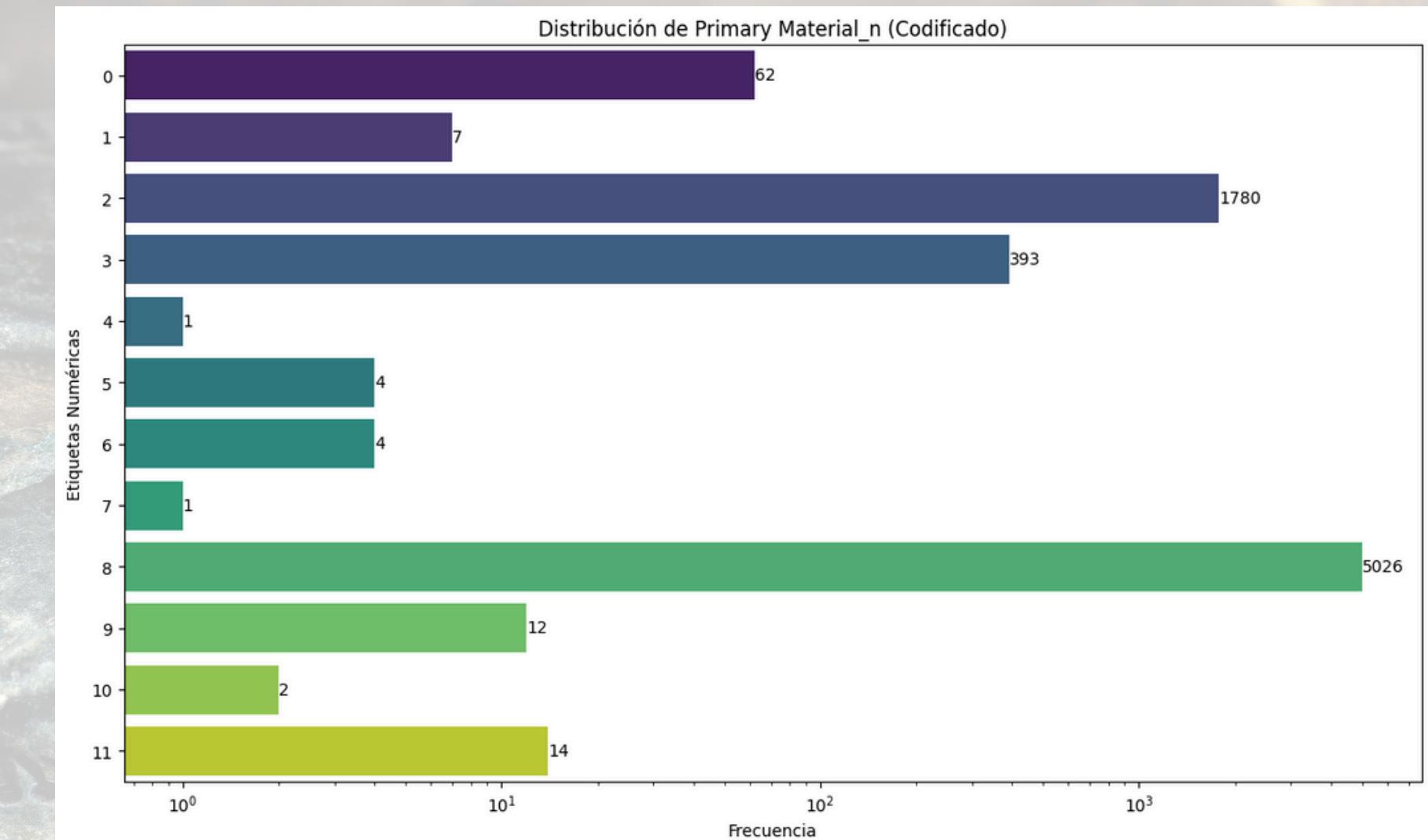
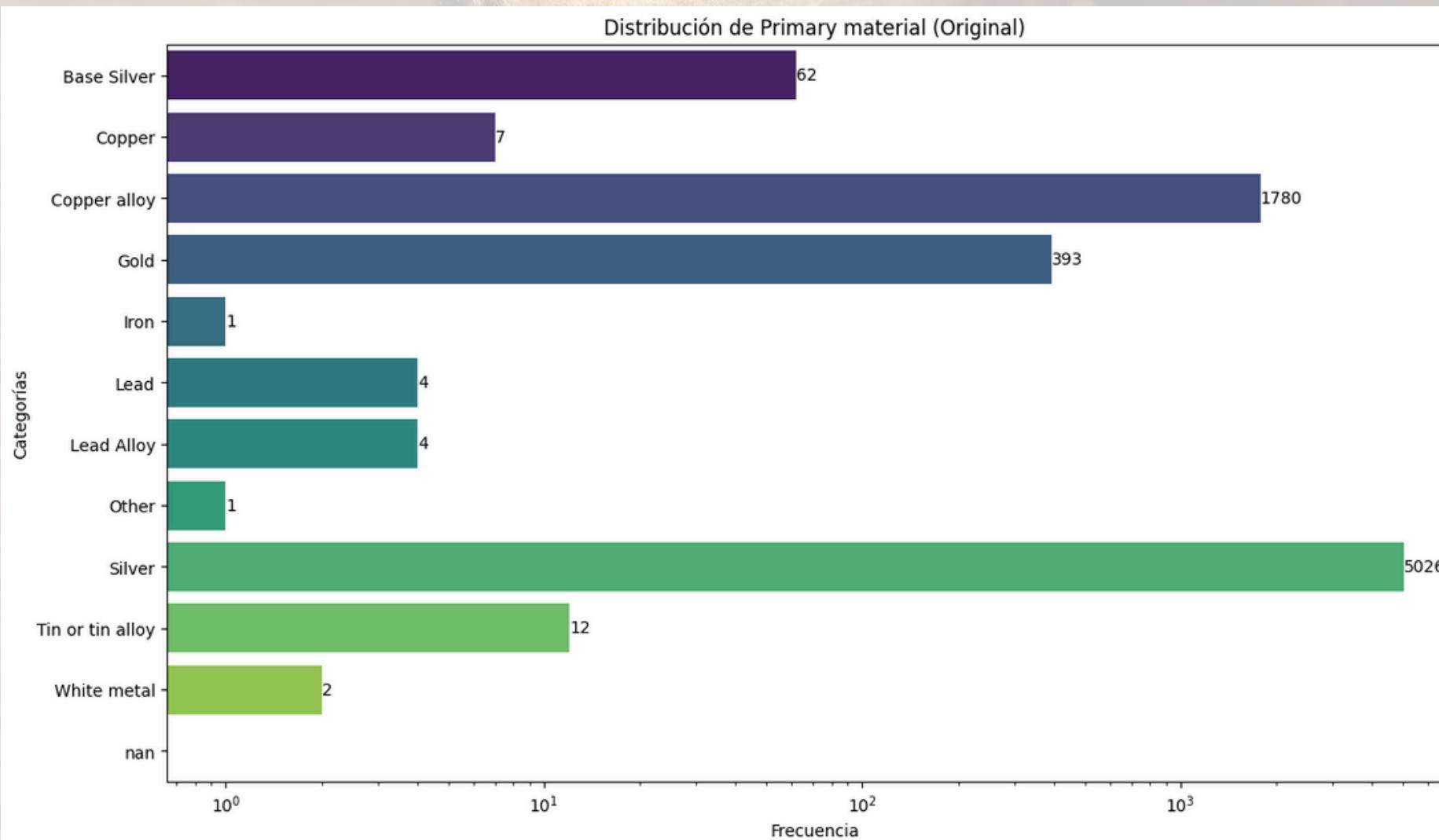




# Procesamiento de los datos

## 2. Convertir las características categóricas en numéricas.

### 2.1. Materia prima.

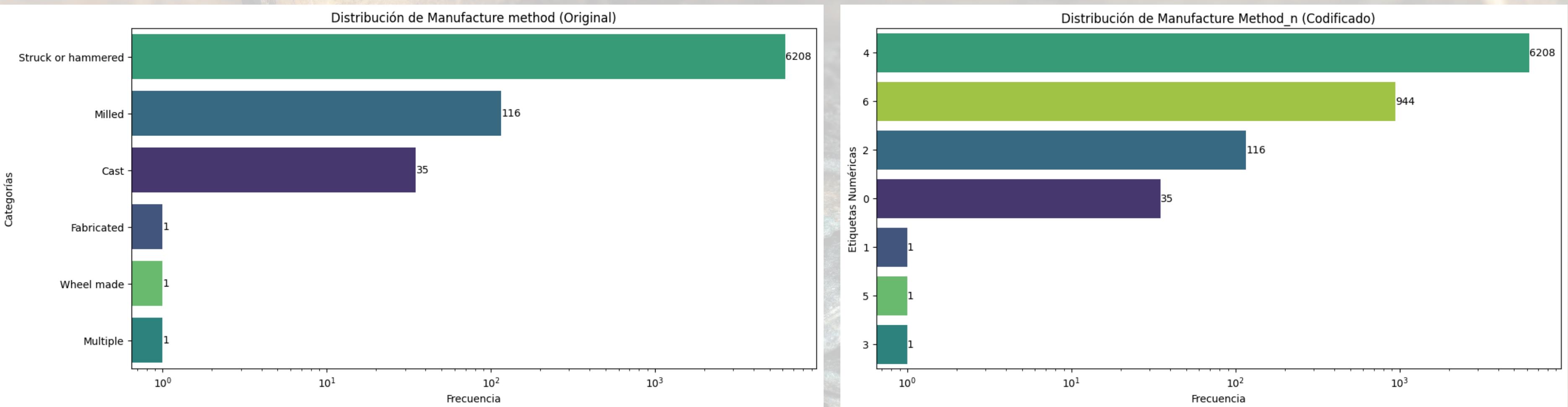




# Procesamiento de los datos

2. Convertir las características categóricas en numéricas.

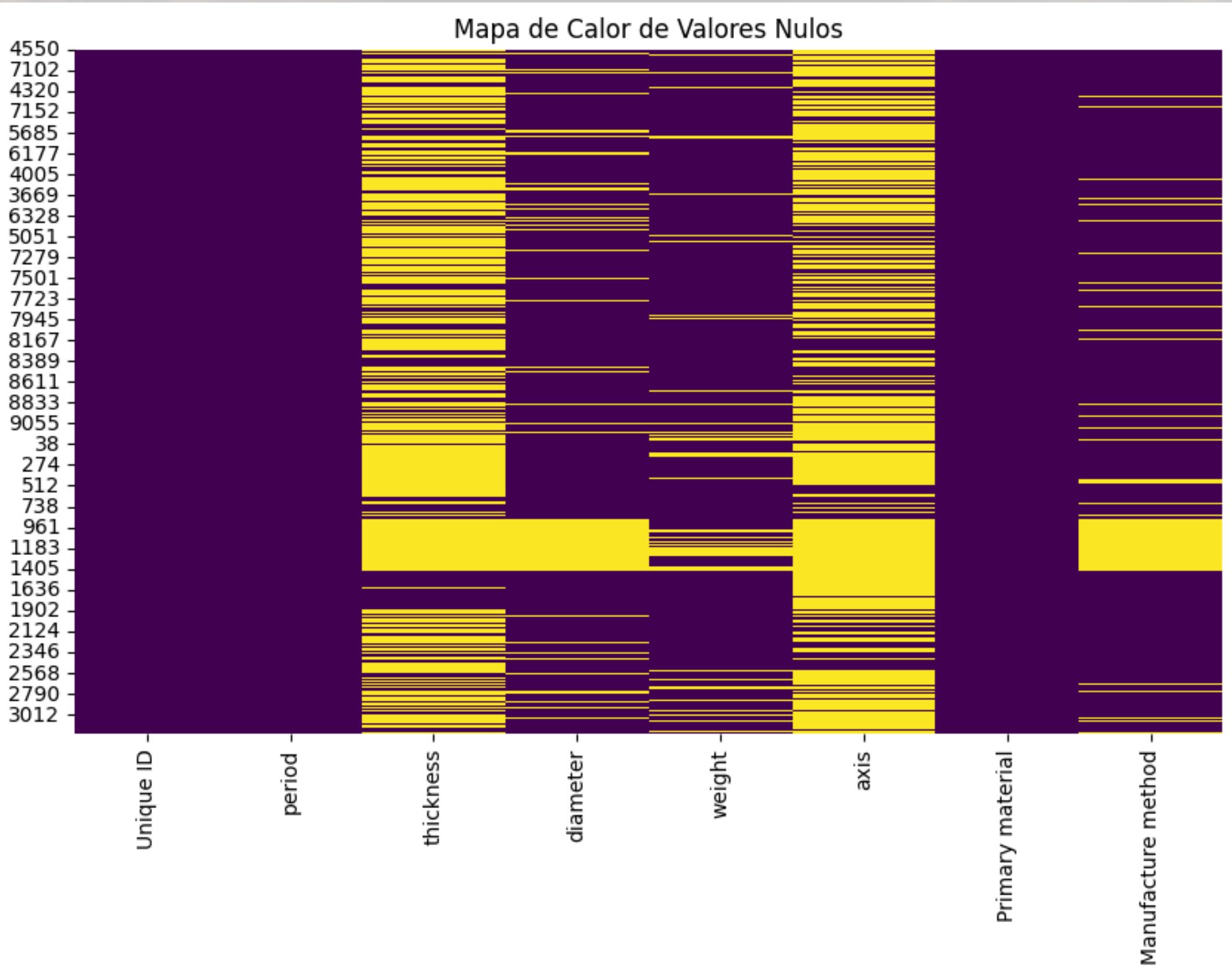
2.2. Método de fabricación.





# Procesamiento de los datos

## 3. Tratar los nulos.



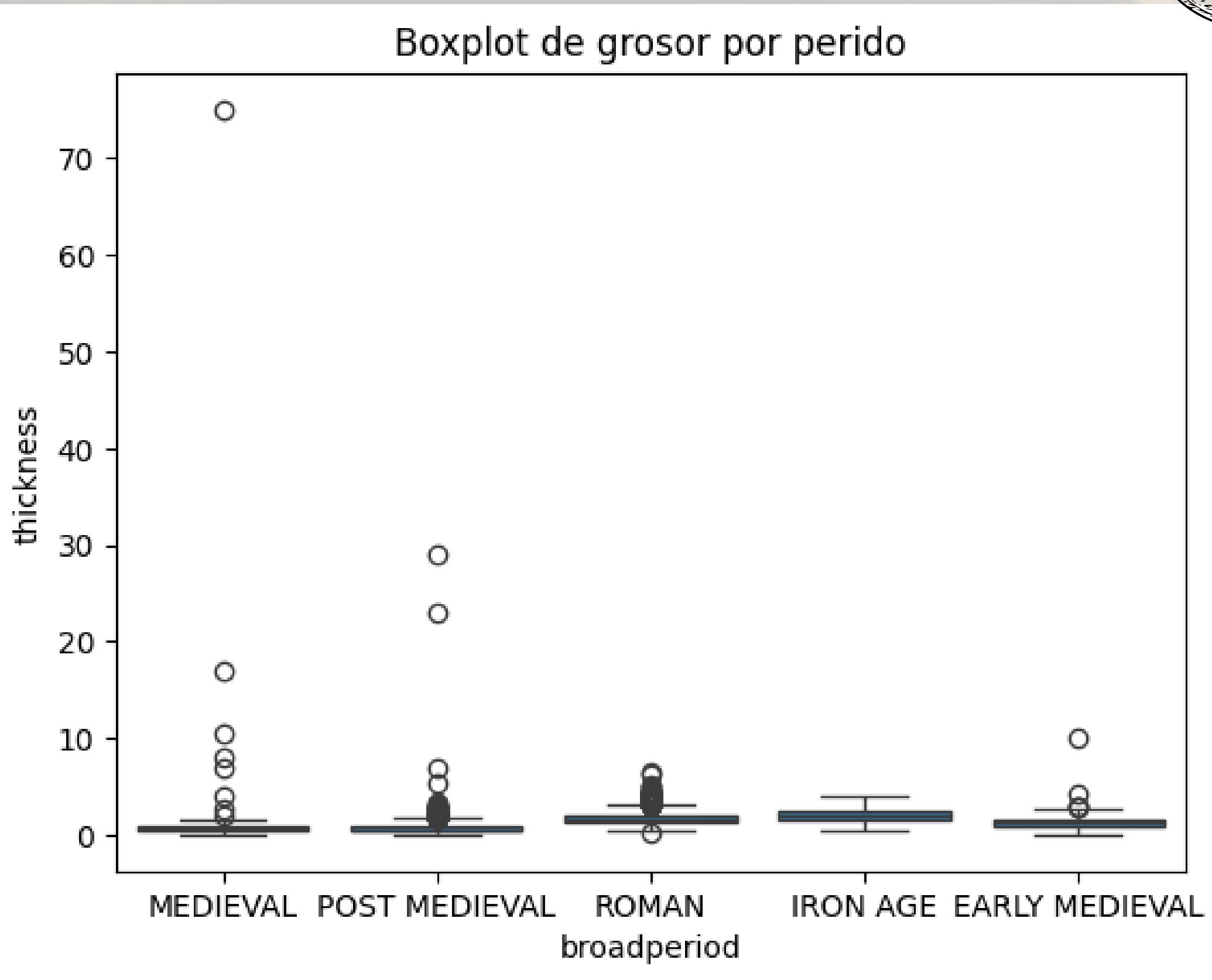


# Procesamiento de los datos



## 3. Tratar los nulos.

Calculo la media, la moda y  
la mediana.



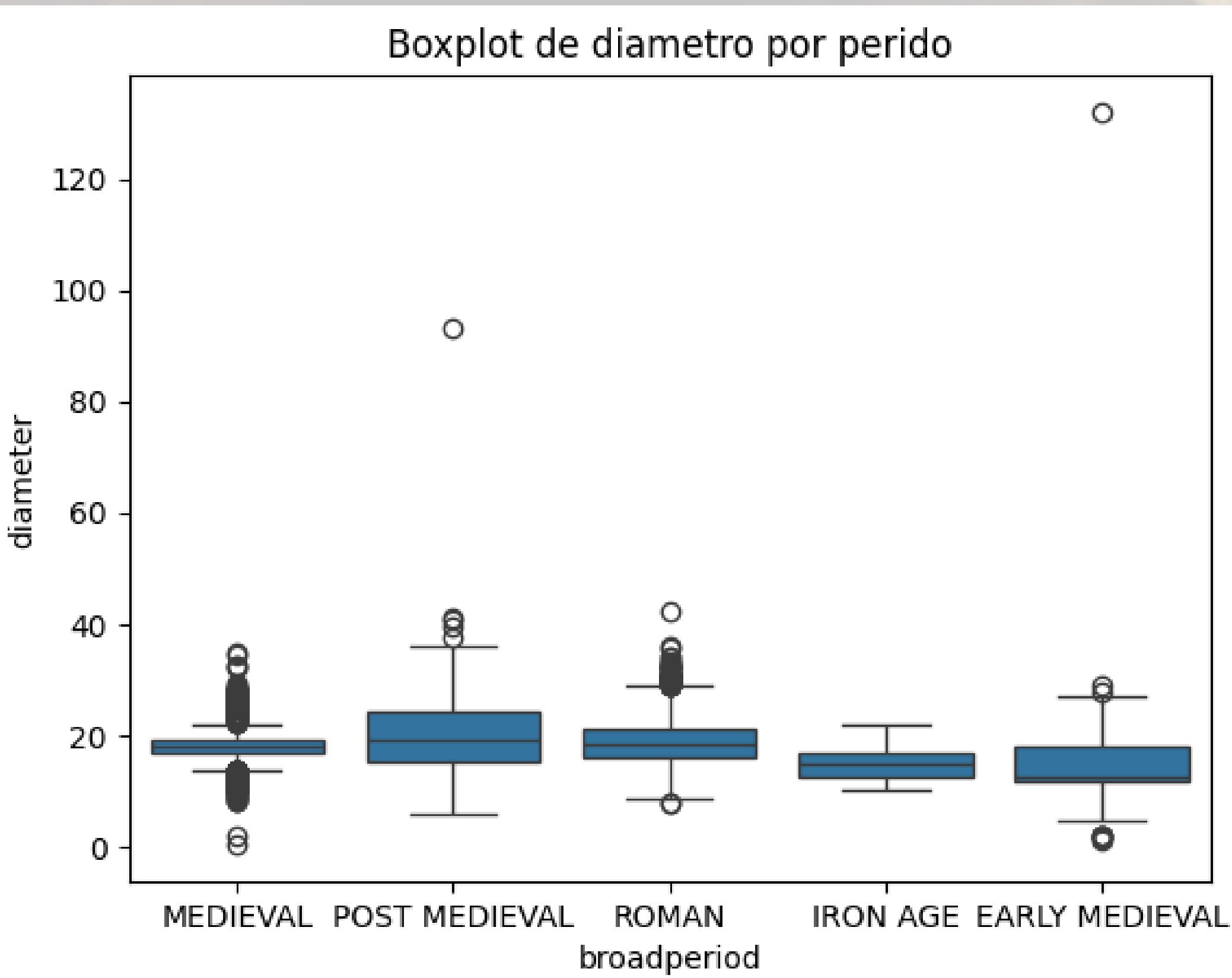


# Procesamiento de los datos



## 3. Tratar los nulos.

Calculo la media, la moda y  
la mediana.



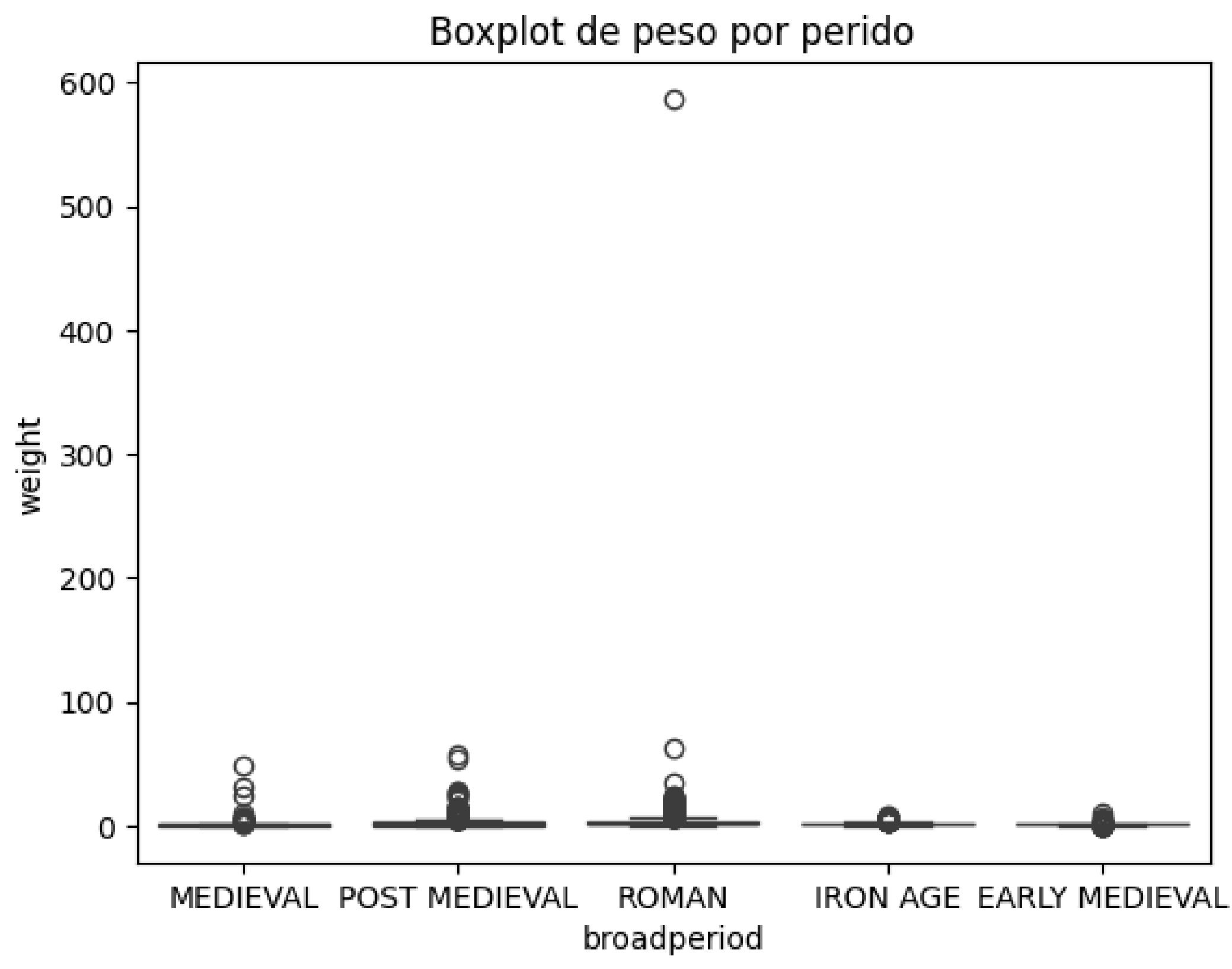


# Procesamiento de los datos



## 3. Tratar los nulos.

Calculo la media, la moda y  
la mediana.



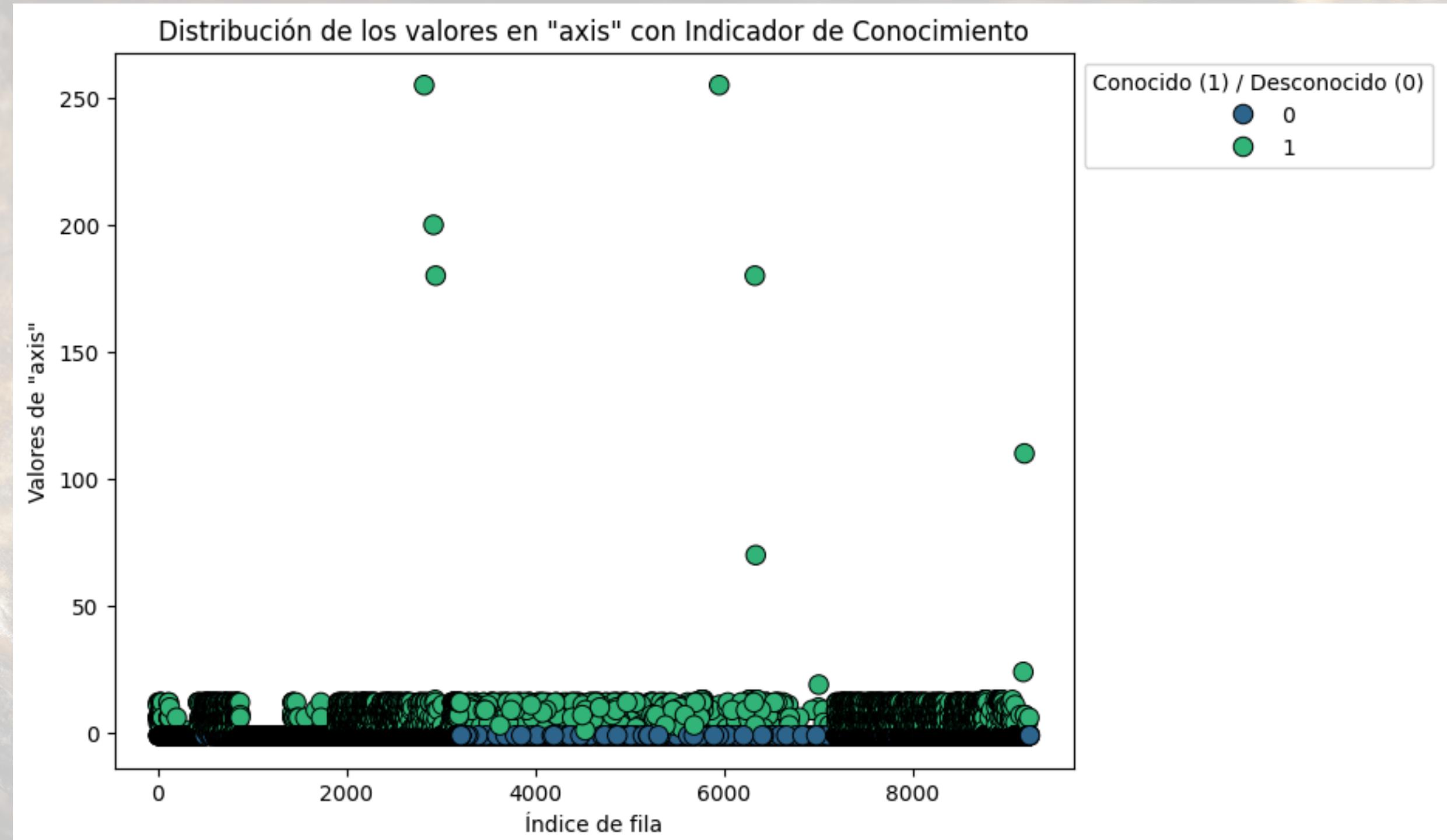


# Procesamiento de los datos



## 3. Tratar los nulos.

Creo un valor fuera de rango (-1) y a través de una columna establecemos si era un valor conocido o desconocido.



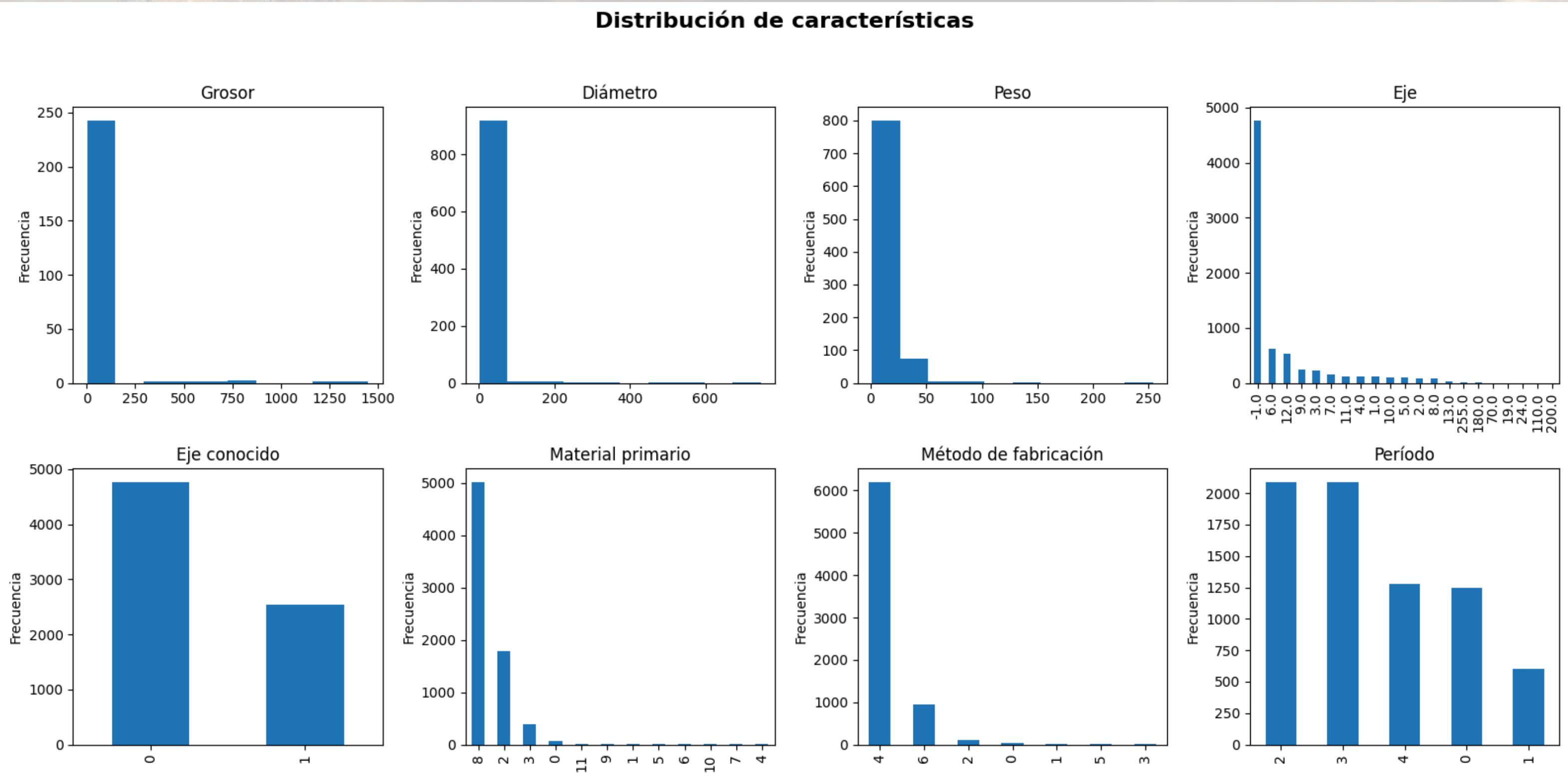
# Modelos

- 5 Modelos supervisados:
  - HistGradient Boosting Classifier.
  - Random Forest Classifier.
  - SVC.
  - Redes Neuronales.
  - Redes Convolucionales.
- 1 Modelo no supervisado:
  - PCA - Random Forest.



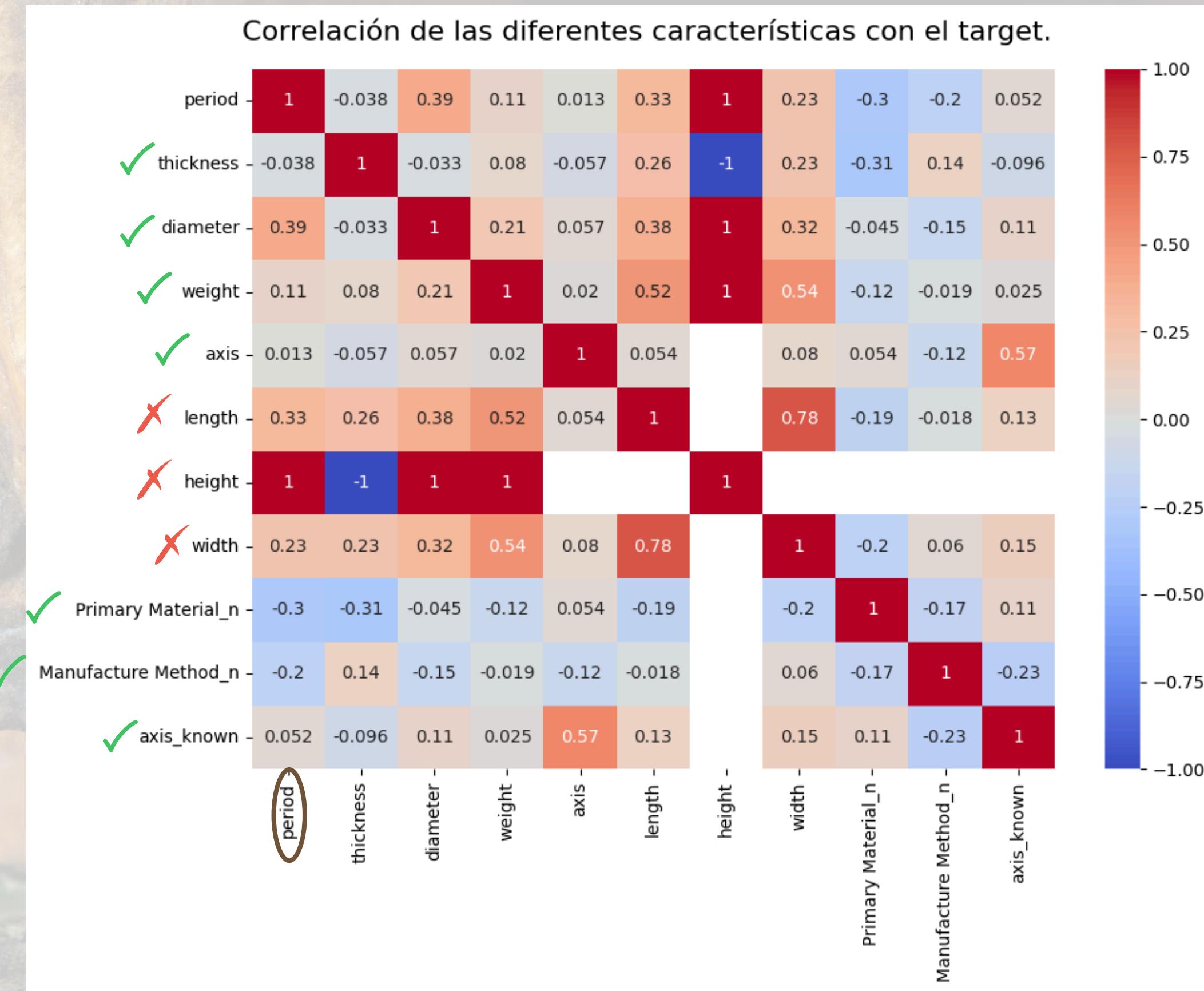
# Elección del target y de las features

**Distribución de características**



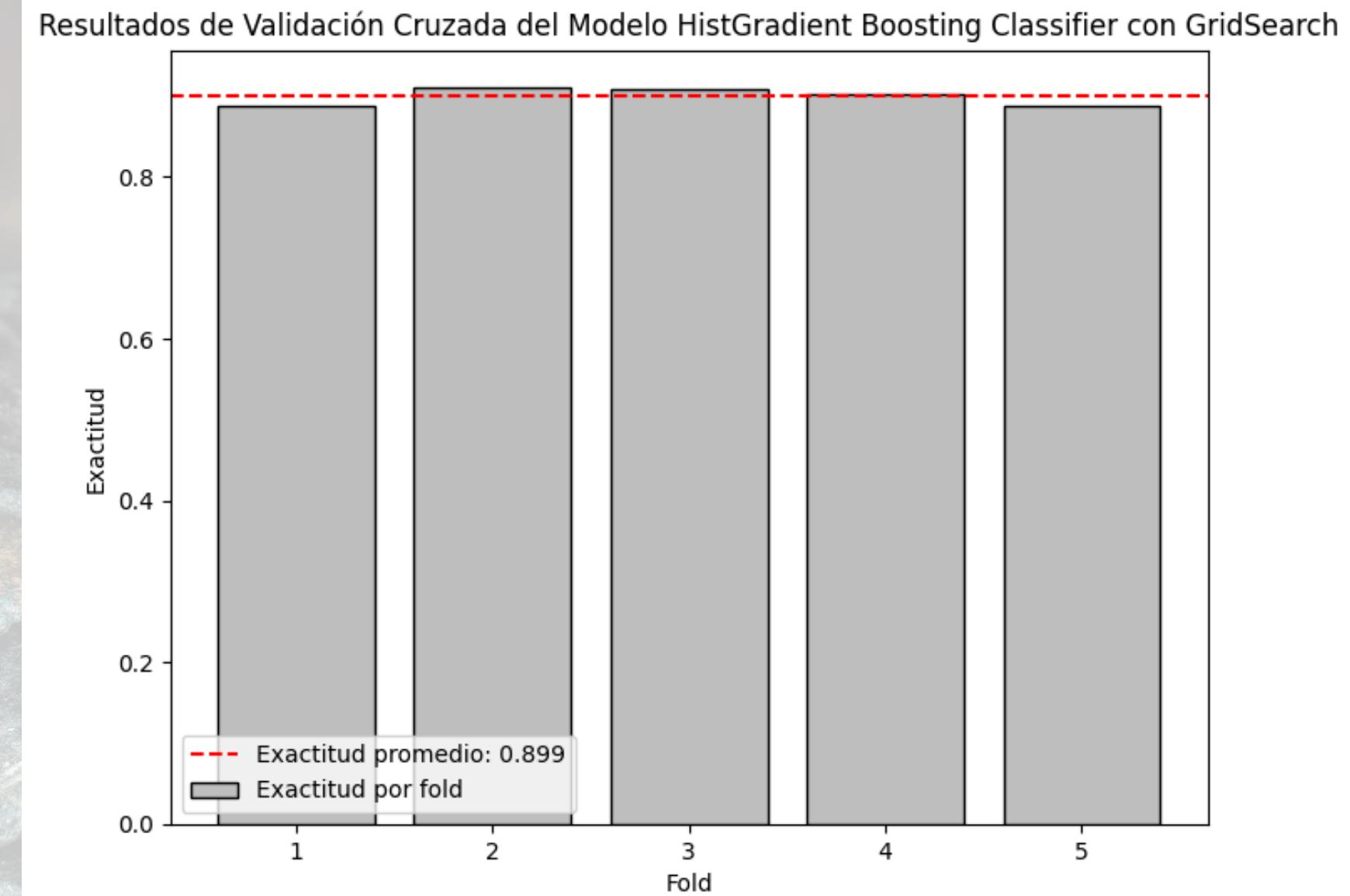
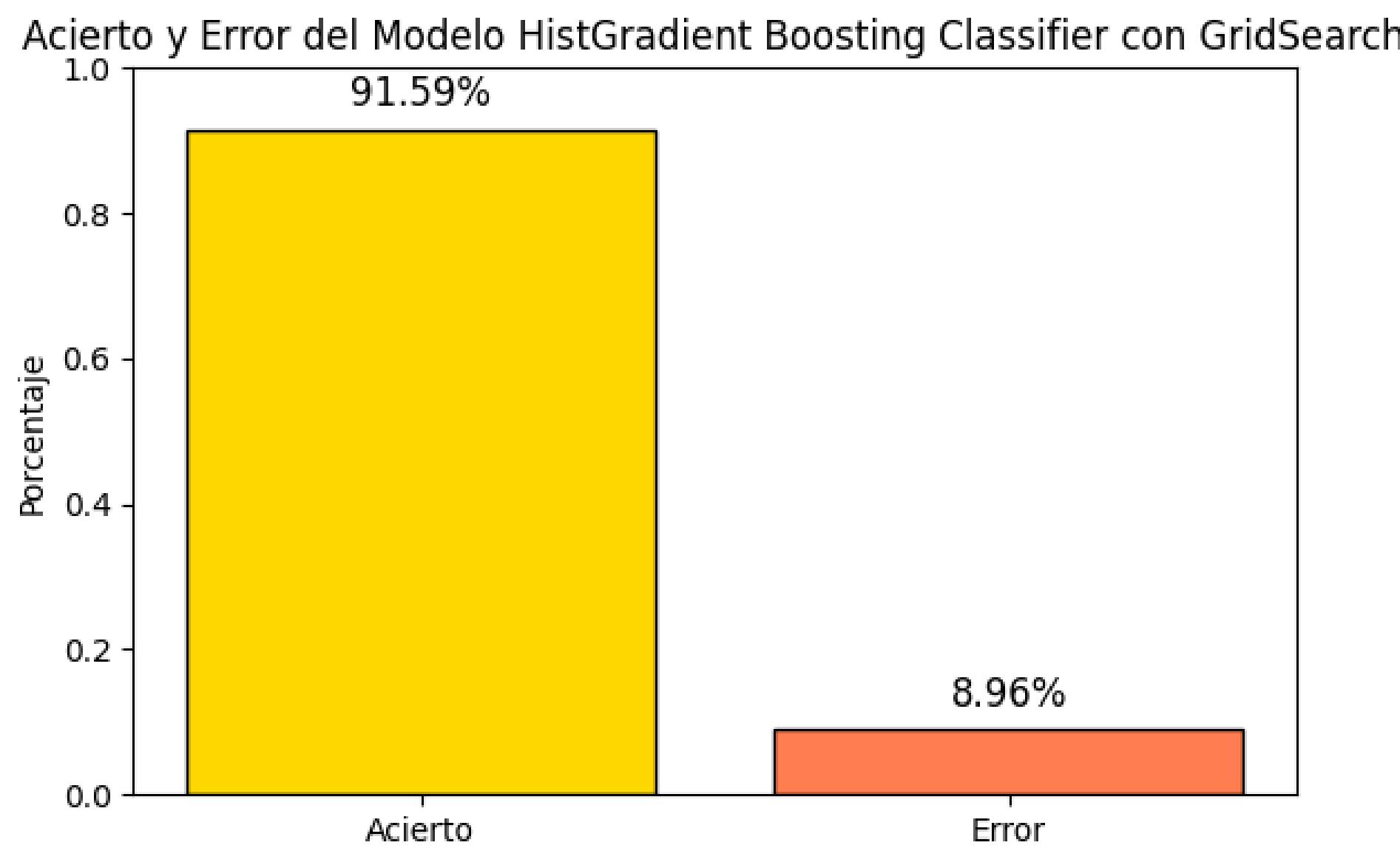


# Elección del target y de las features



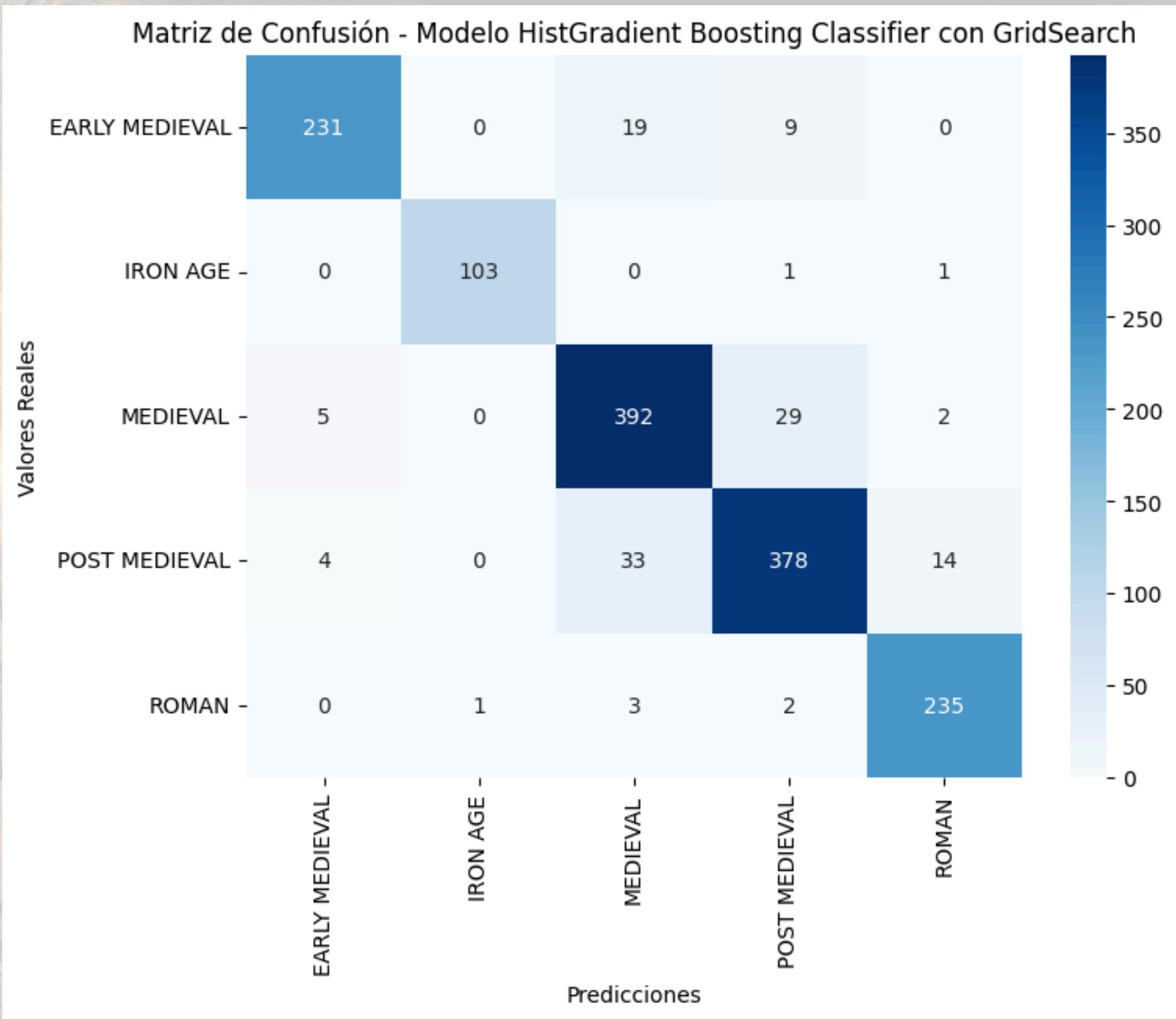


# HistGradient Boosting Classifier.



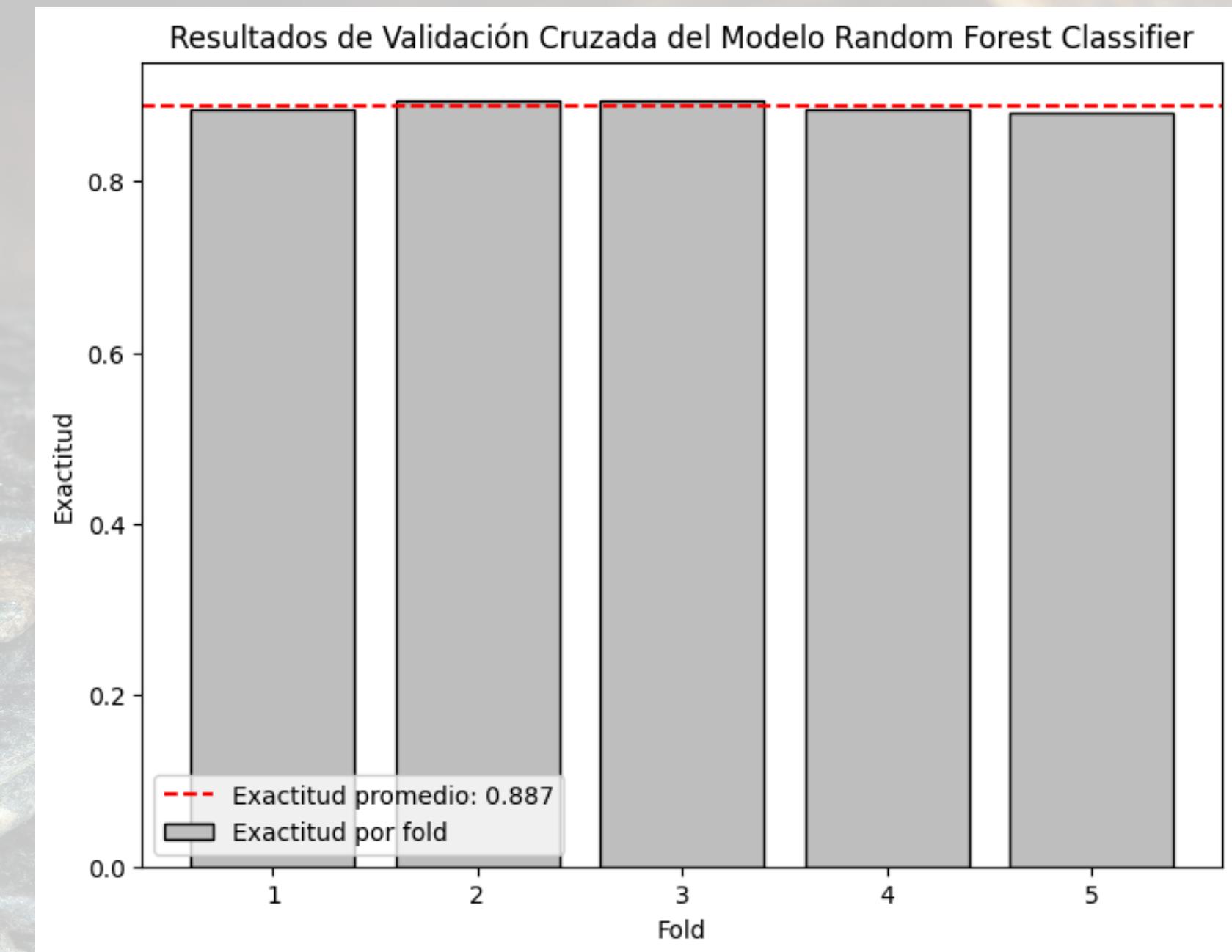
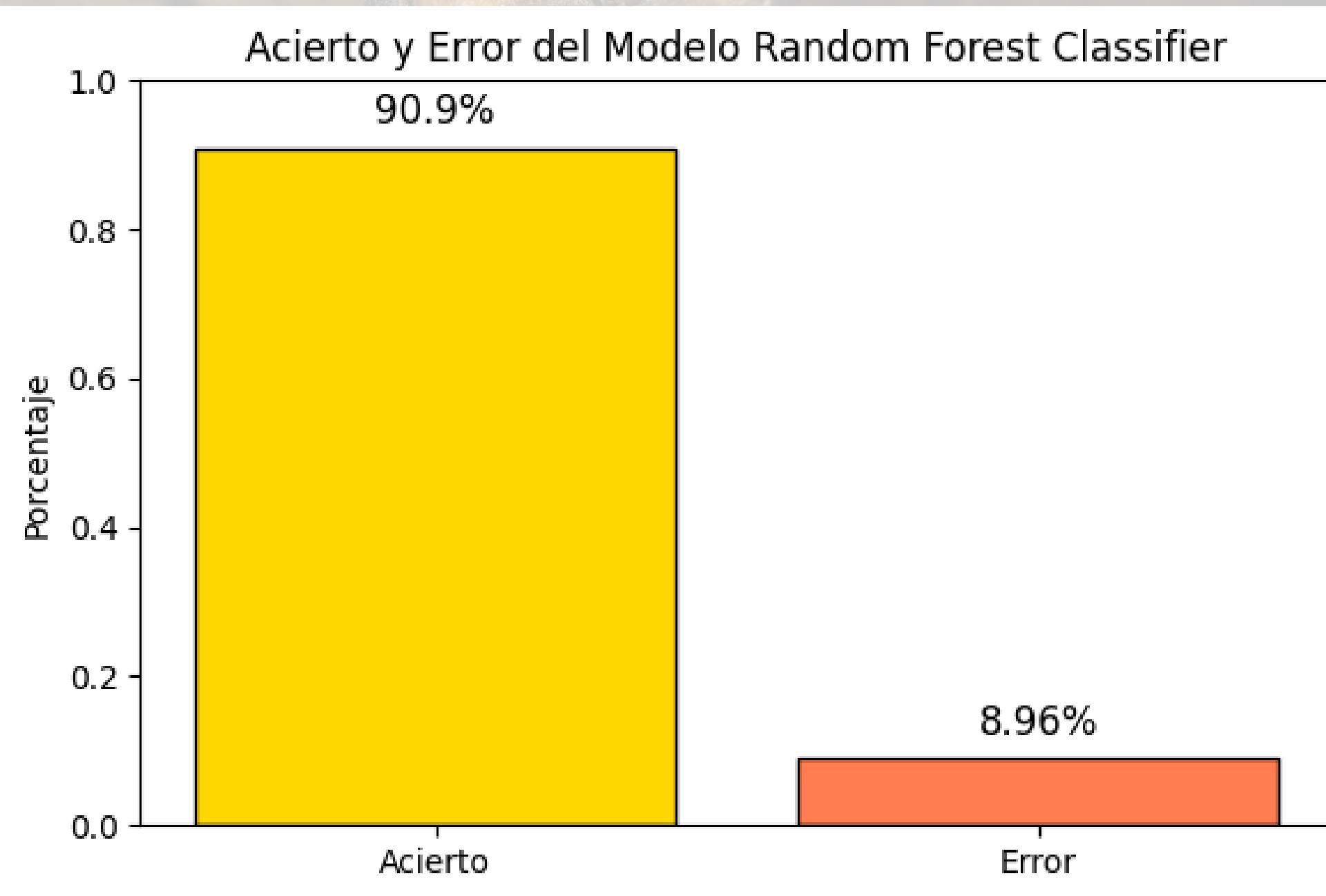


# HistGradient Boosting Classifier.



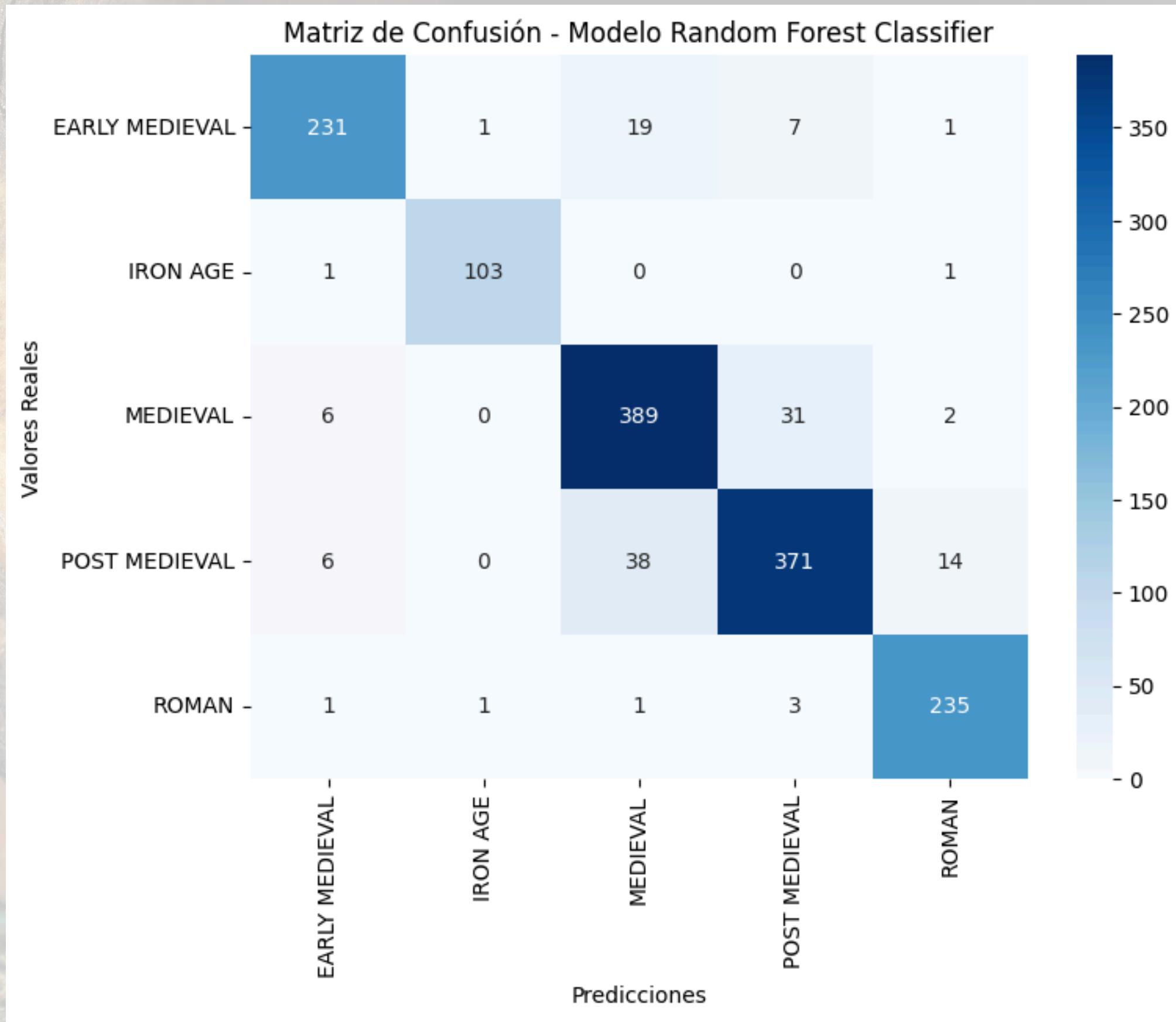


# Random Forest Classifier.





# Random Forest Classifier.

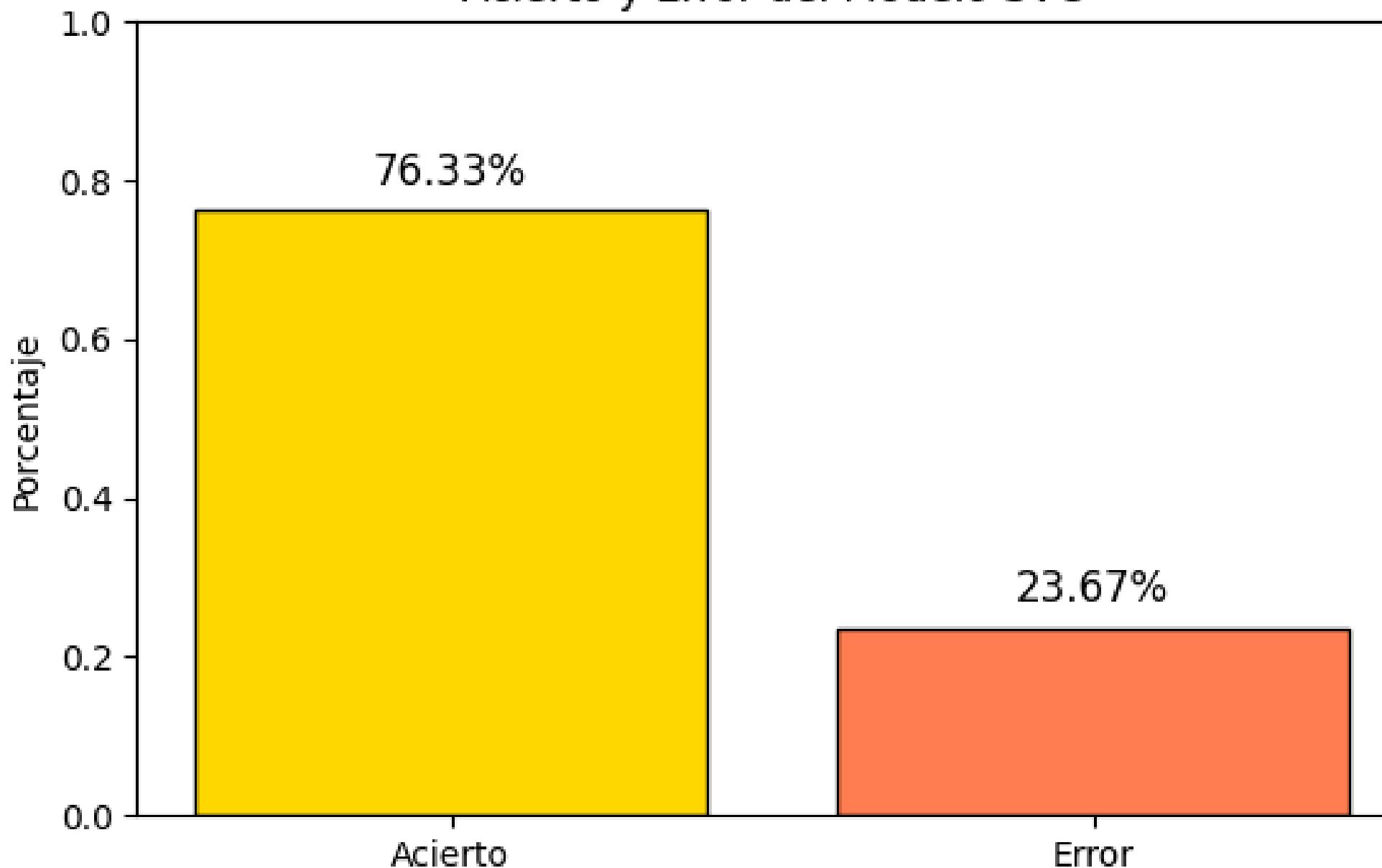




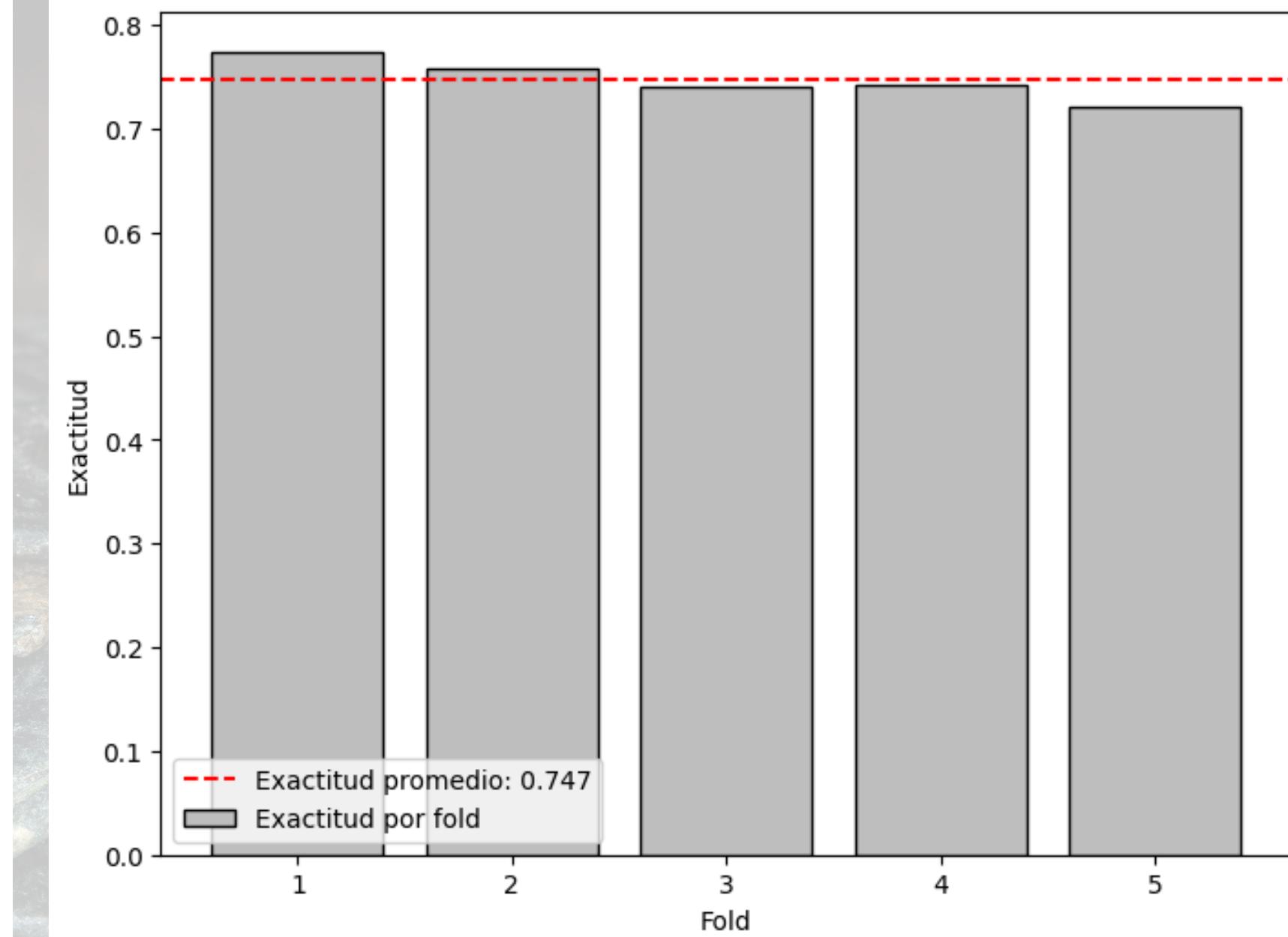
sVC



Acierto y Error del Modelo SVC



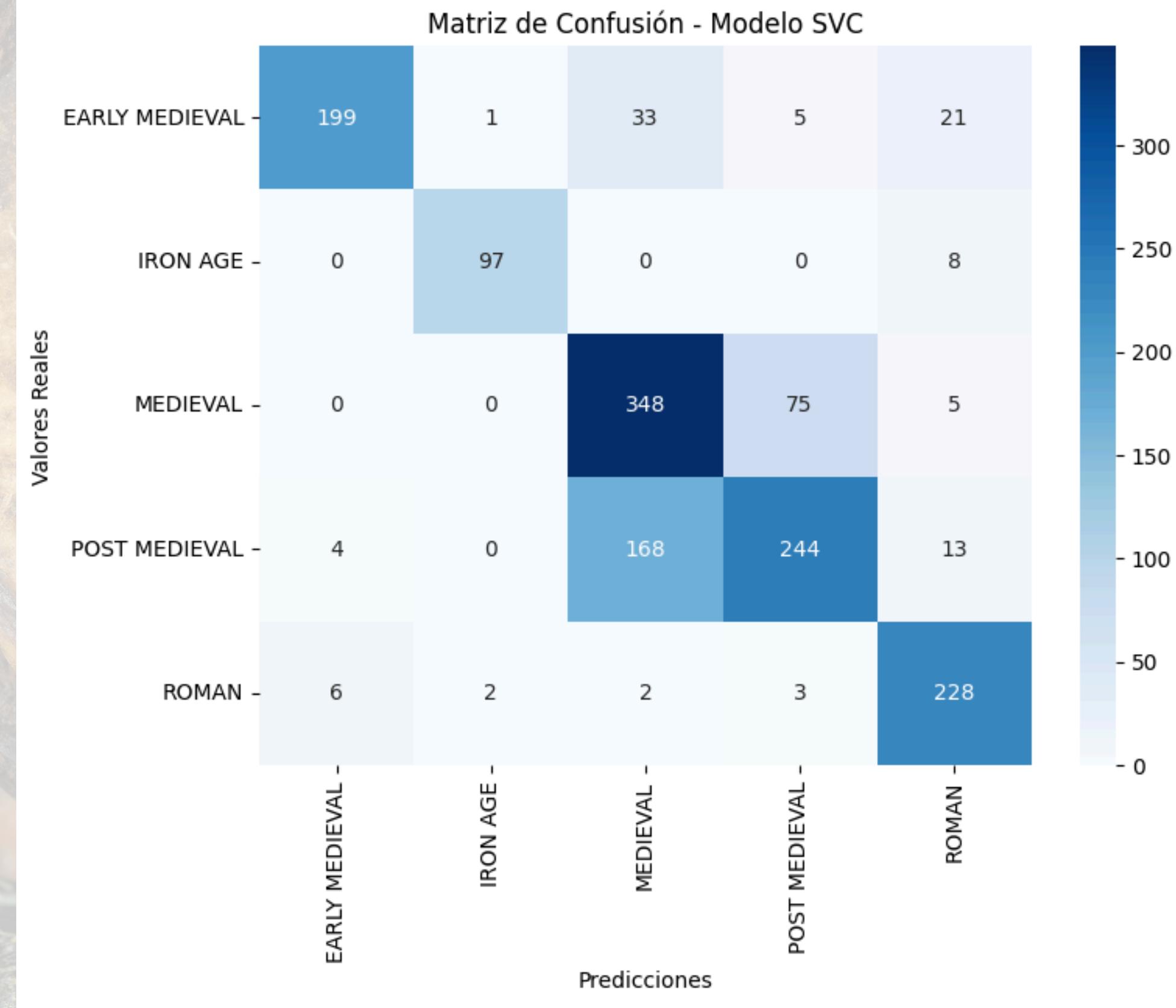
Resultados de Validación Cruzada del Modelo SVC



# SVC



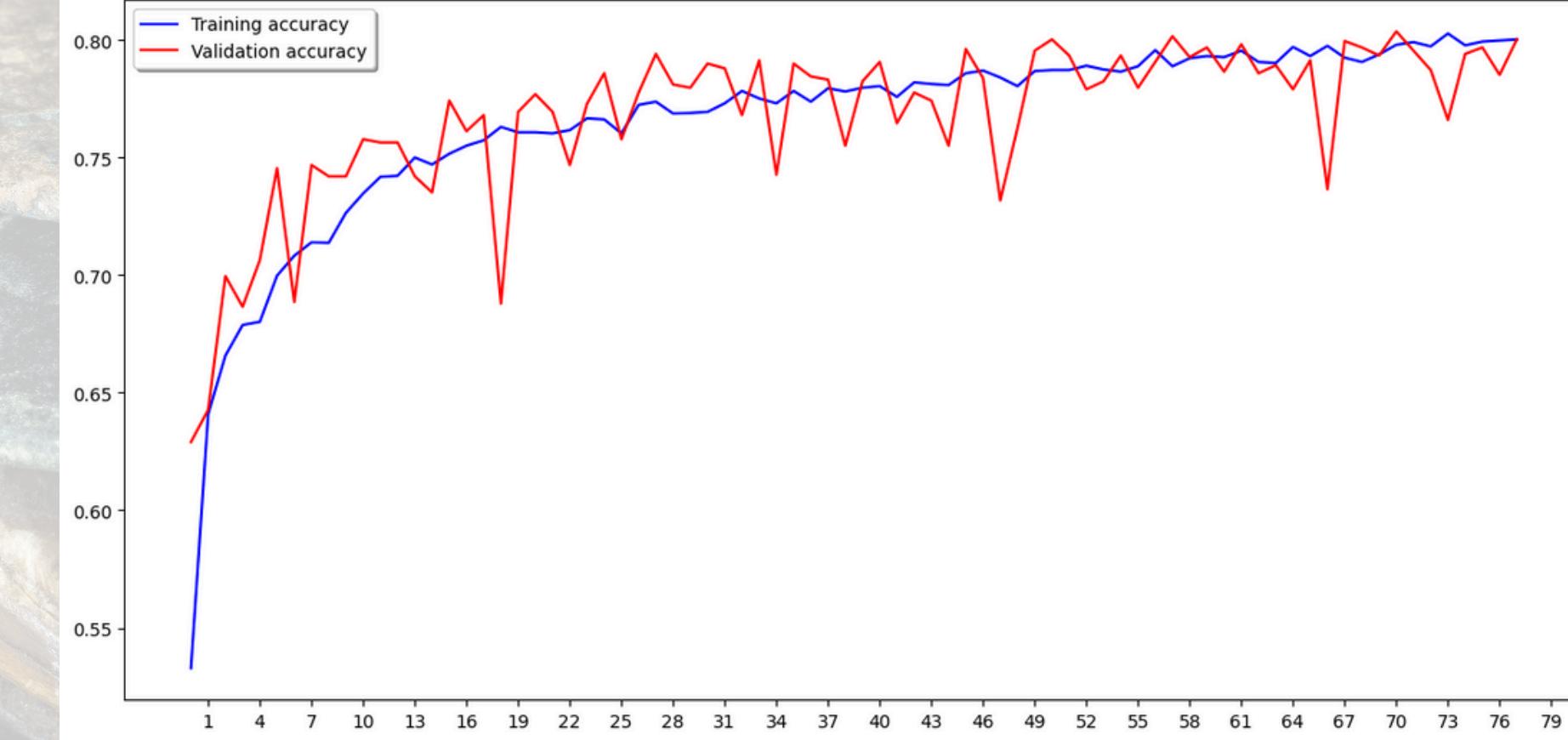
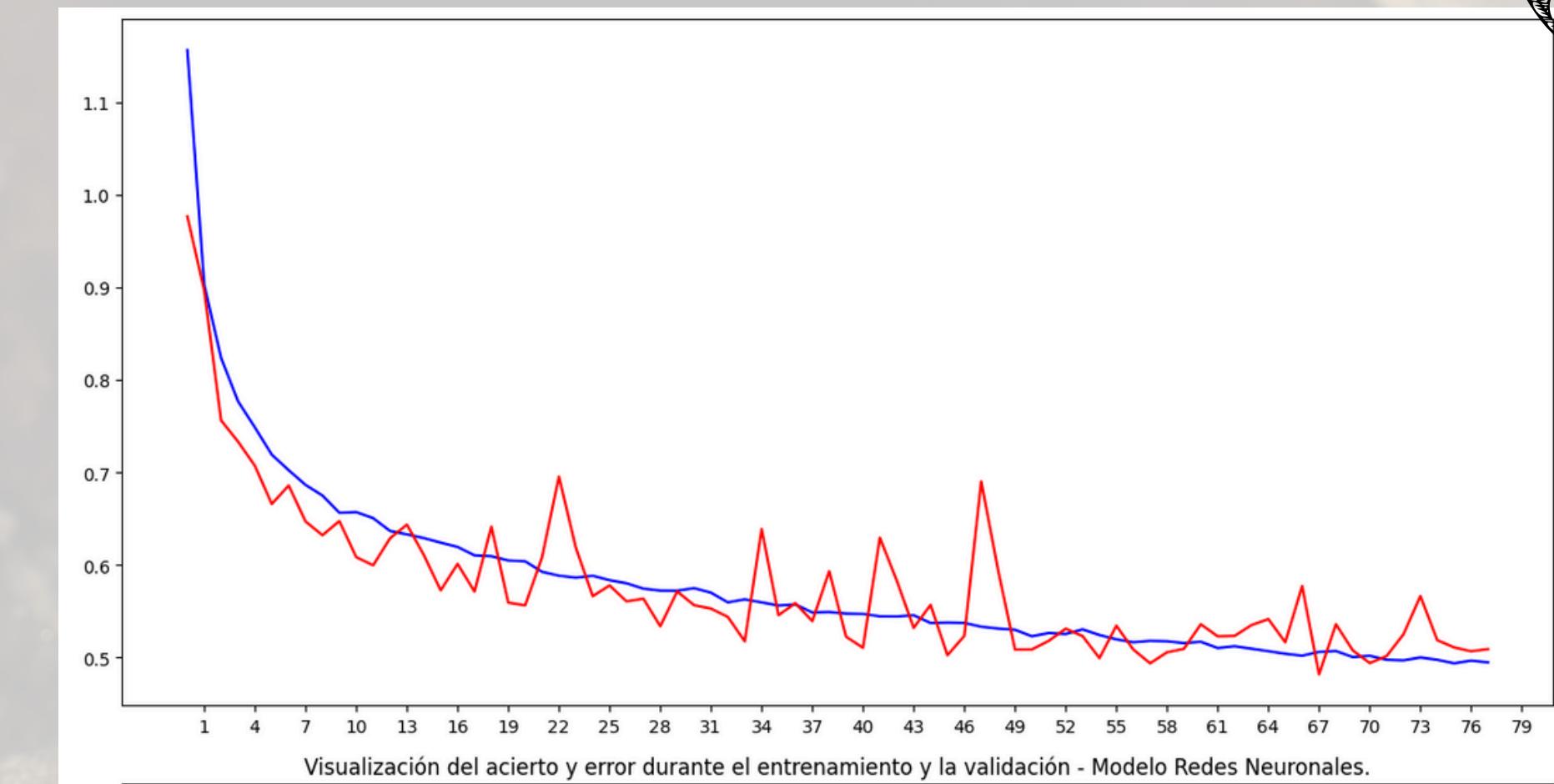
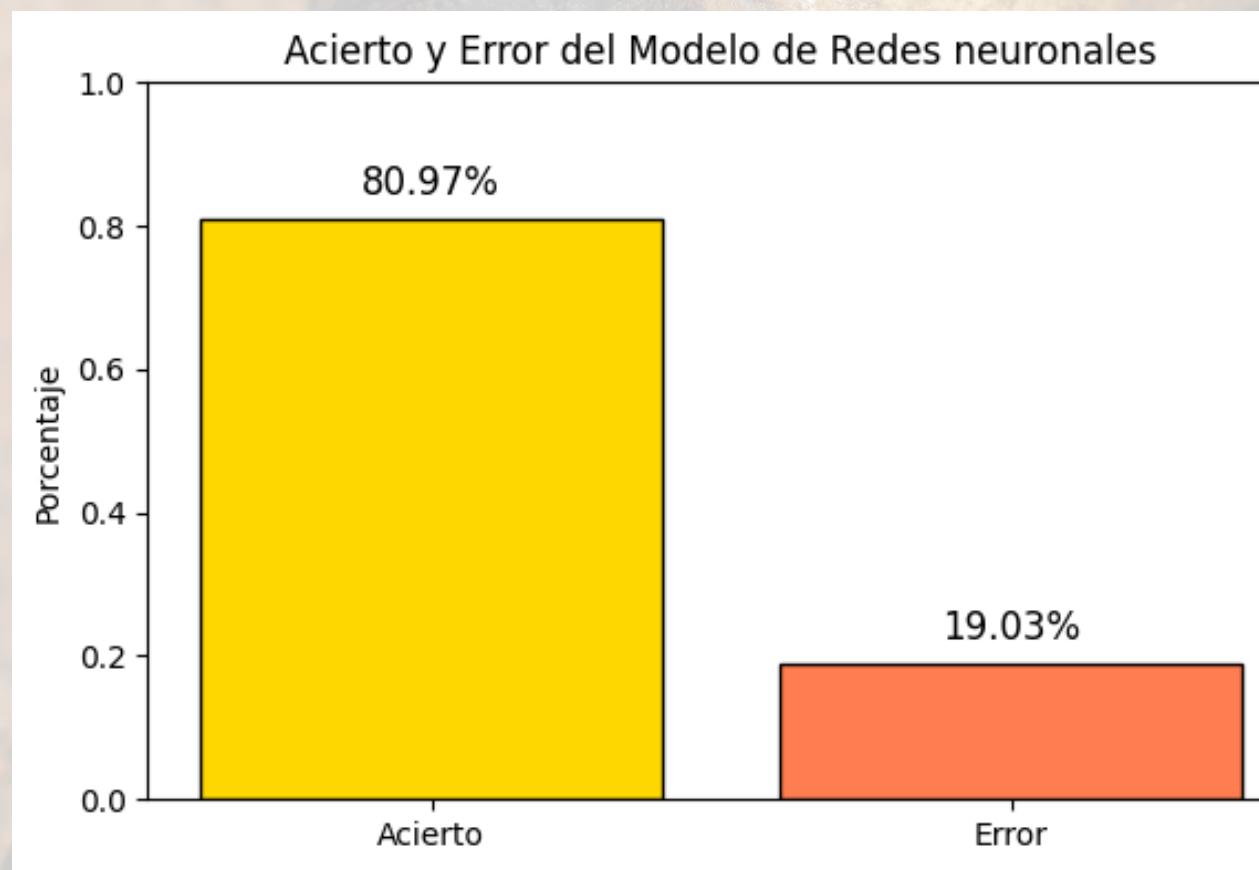
Matriz de Confusión - Modelo SVC



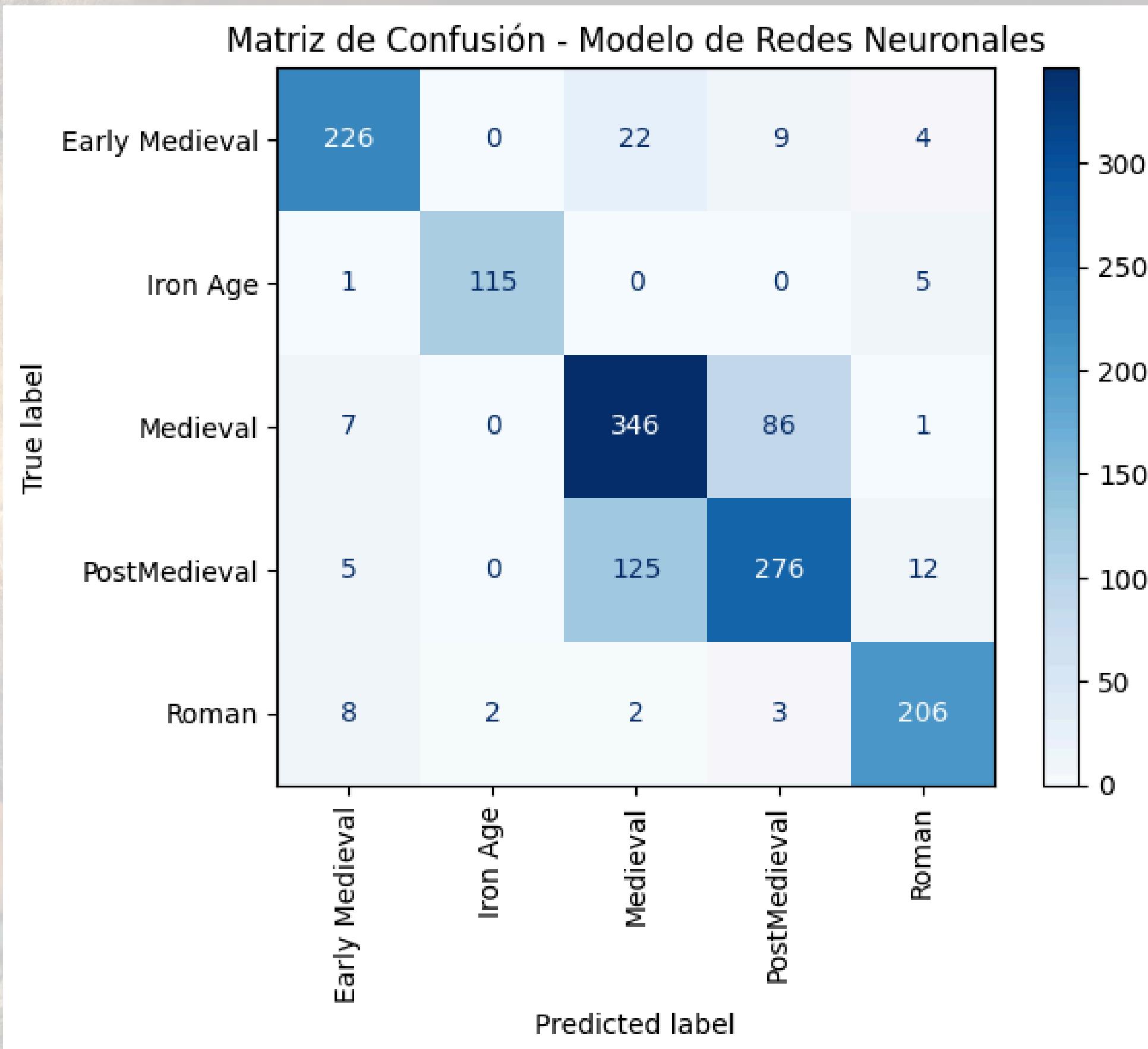


# Redes neuronales

- 4 Capas
- Activador: 'relu'
- Salida 'softmax'



# Redes neuronales

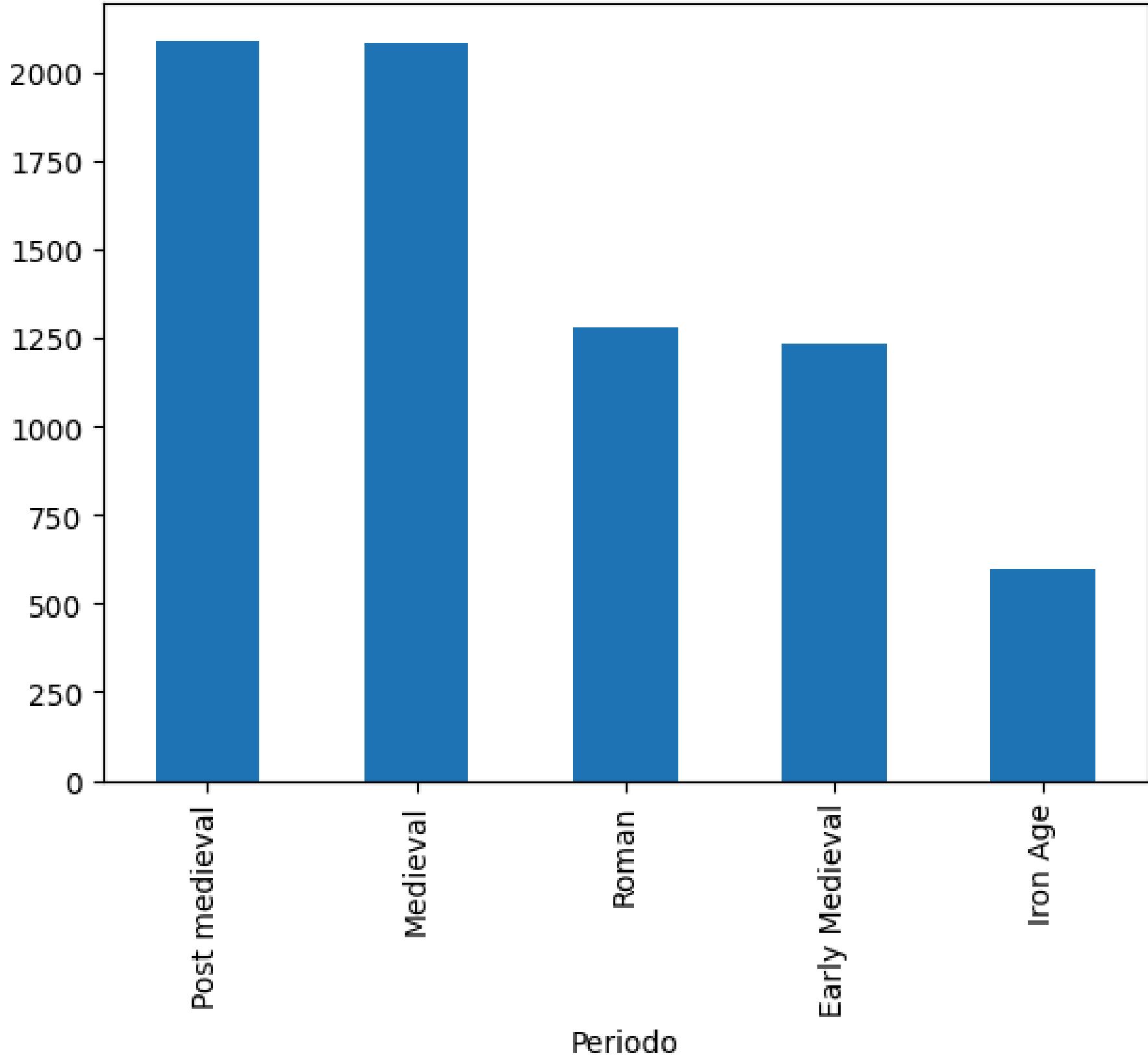




# Redes convolucionales



Distribución de las imágenes según la época.





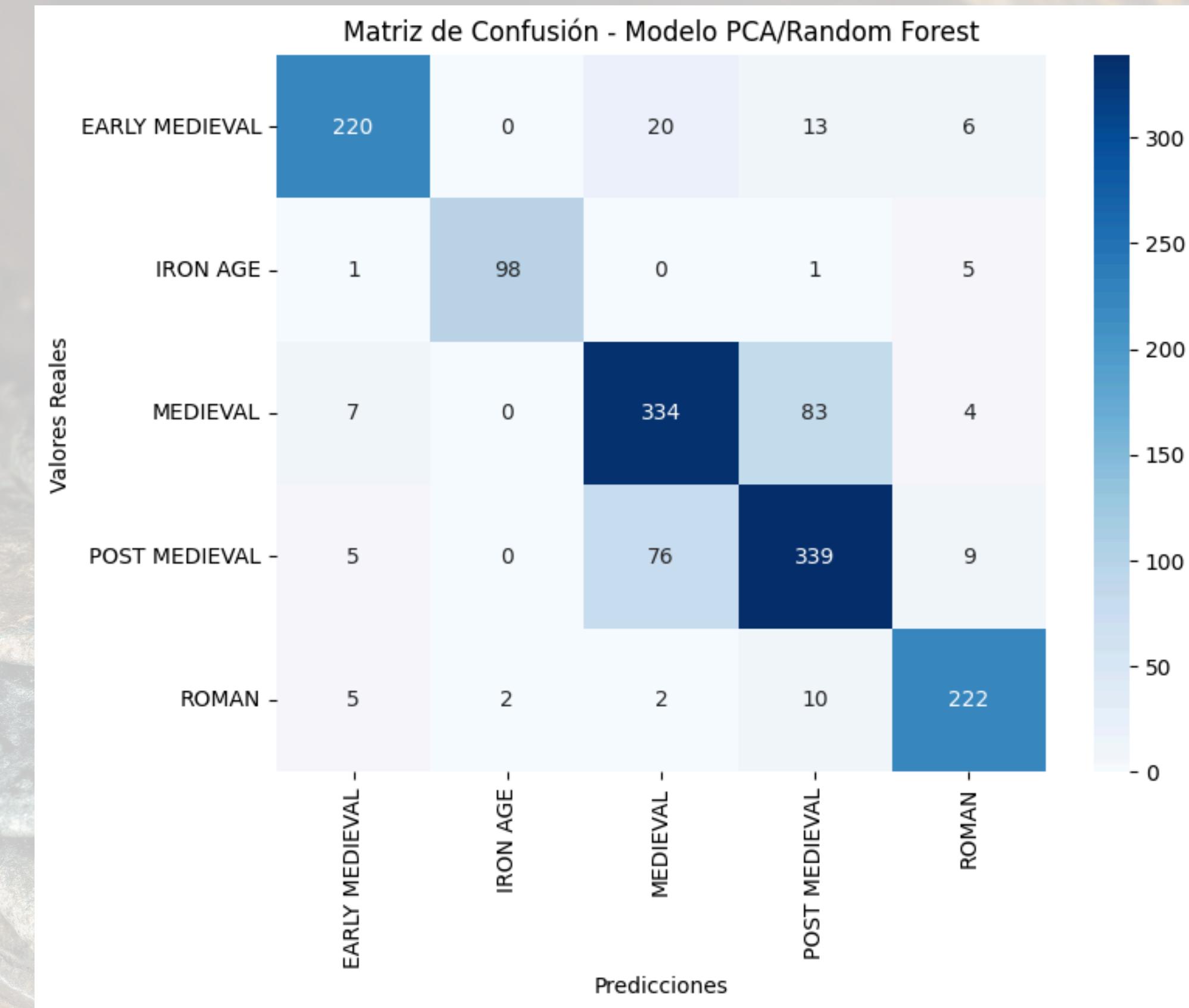
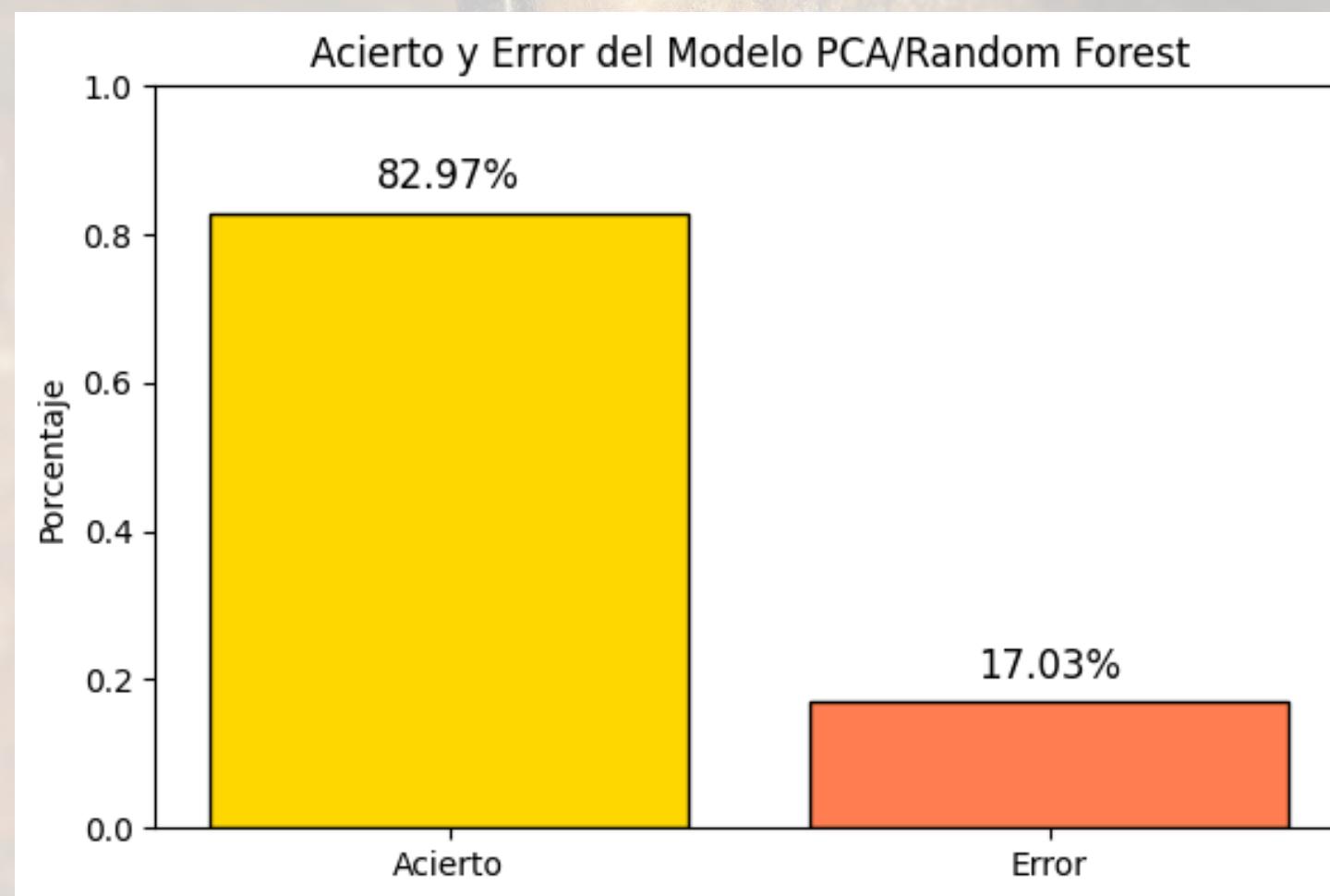
# Redes convolucionales

Matriz de Confusión - Modelo Redes convolucionales

Verdaderas	Predicciones	Matriz de Confusión - Modelo Redes convolucionales				
		Early Medieval	Iron Age	Medieval	Post Medieval	Roman
Early Medieval	130	3	76	46	7	
Iron Age	4	92	3	2	2	
Medieval	50	6	266	85	18	
Post Medieval	22	10	75	263	23	
Roman	15	3	20	20	215	

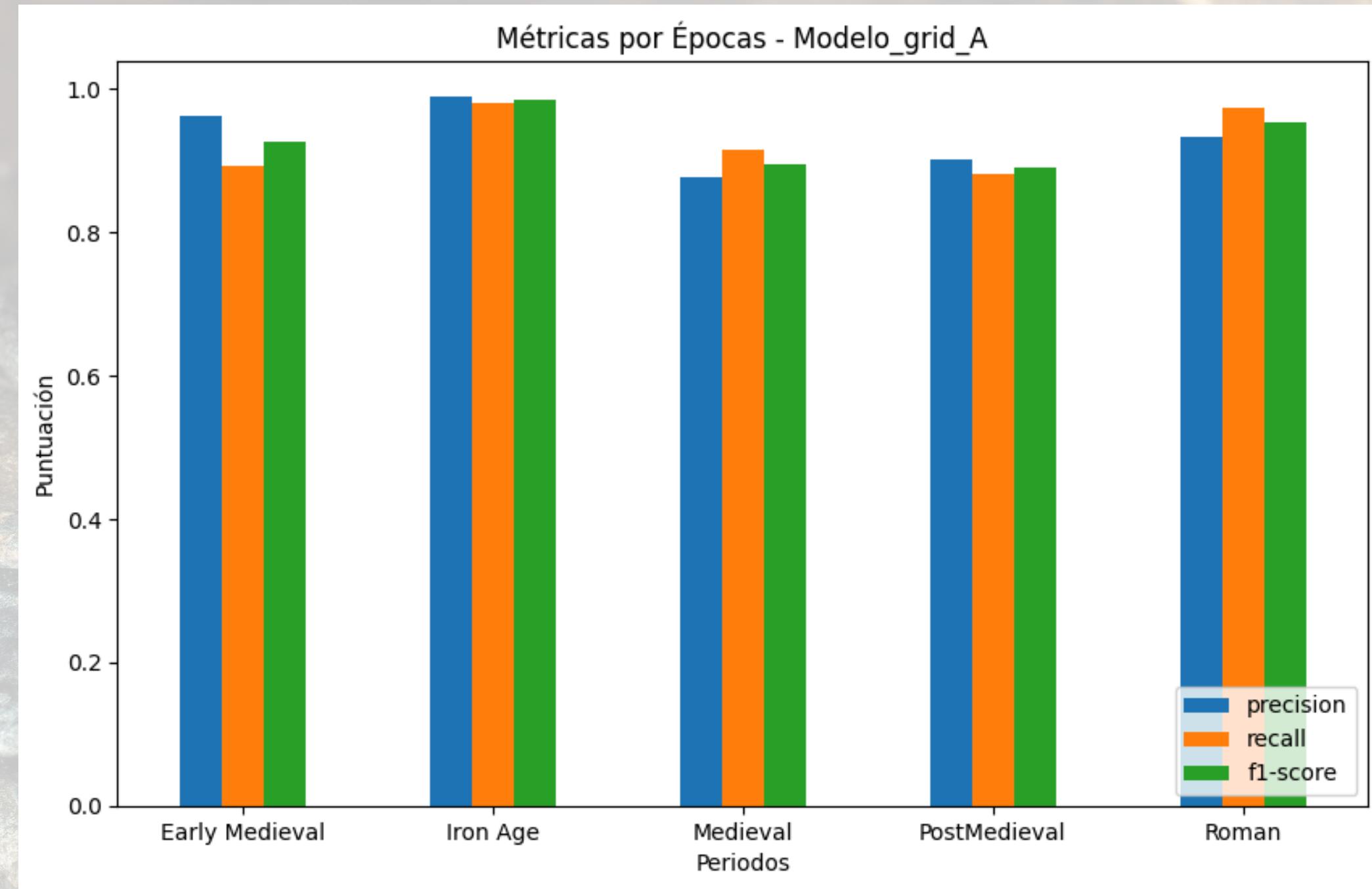
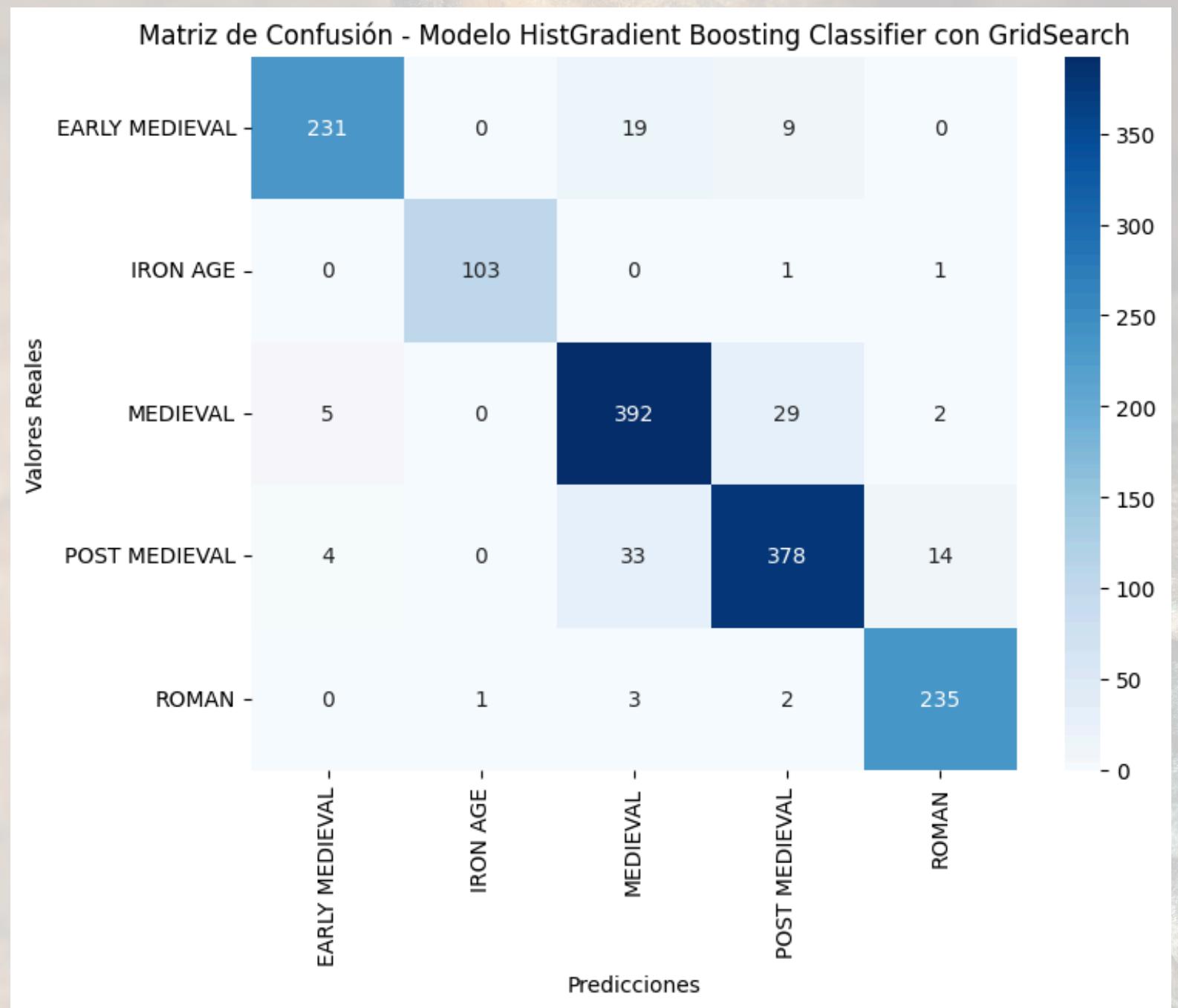


# PCA - Random Forest



# Modelo elegido

## HistGradient Boosting Classifier.



# Escalabilidad

- Los datos provienen del Museo Británico y su procedencia en su mayoría es de origen británico.
  - Nutrir los datos introduciendo nuevos registros procedentes de otros países.
  - Generalizará mejor.

# Escalabilidad

- Introducir nuevas features, por ejemplo, reconocimiento de las inscripciones de los anversos o en los reversos.
  - Va a permitir una mejor identificación de la iconografía y la epigrafía.
  - Generalizará mejor y será más preciso.



GRACIAS



# Relicia

## Clasificación numismática

Vanessa López Aguilera  
Febrero 2025