

# Anexos

Como anexos se presentan los diferentes scripts que se emplearon en la implementación del algoritmo en su totalidad, estos scripts se presentan en un repositorio de GitHub, se presenta un instructivo para la correcta ejecución de los scripts y de los datos a usar.

[Link GitHub](#)





## **BLR\_imputation\_perovskite**

Mediante este proyecto se implementa un algoritmo que realiza dos procesos de imputación, los modelos emplean el criterio BLR, primero se realiza imputación para las variables de entrada y salida, y como segundo modelo se realiza imputación mediante estimación para las variables de rendimiento/salida: *Voc*, *Jsc*, *PCE* y *FF*.

Este proyecto se hizo con la intención de ser replicable y utilizable sin importar la base de datos a utilizar, aclarando que algunas etapas, como lo es la de preprocesamiento, son únicas en este proyecto, dado que solucionan problemas y extraen características únicas de las bases de datos de interés presentadas.



## **Instrucciones para los scripts**

Para el correcto funcionamiento de este proyecto es importante seguir las instrucciones y tener en cuenta las pautas mencionadas. Cada script se encuentra relacionado con una etapa en específico.

-  **Preprocesamiento:** Esta etapa se encuentra relacionada con los scripts "Preprocesamiento\_base\_datos\_imputar" y "Preprocesamiento\_base\_datos\_complementaria".
-  **Construcción y entrenamiento del modelo:** Esta etapa se encuentra relacionada con los scripts "Proyecto\_de\_grado\_imputacion" y "Proyecto\_de\_grado\_estimacion".
-  **Evaluación:** Esta etapa se encuentra relacionada con los scripts "Evaluacion\_imputacion" y "Evaluacion\_estimacion".
-  **Modelo:** Esta etapa es el resultado de las etapas anteriores, por medio de este se da un vistazo a los modelos de imputación y de estimación anteriormente mencionados, el script "Imputacion" y el script "Estimacion" se encuentran relacionados a esta etapa.



Además para el correcto funcionamiento de este proyecto es importante la importación o descarga de los datos relacionados a cada script, dado el tamaño de estos se optó por crear un archivo en la nube que permite su descarga Archivos\_proyecto.

El archivo "[Archivos proyecto](#)" cuenta con 3 carpetas:

-  **Bases\_datos:** Corresponden a las bases de datos utilizadas sin preprocesar "base\_datos\_sintesis", "Perovskite\_database\_content\_all\_data" que fungen como entradas de los scripts "Preprocesamiento\_base\_datos\_imputar" y "Preprocesamiento\_base\_datos\_complementaria".
-  **Datos\_preprocesados:** En esta se encuentran la base de datos preprocesada "datos\_sintesis" que funge como base de datos de entrenamiento en los scripts

"Proyecto\_de\_grado\_imputacion" y "Proyecto\_de\_grado\_estimacion", además se encuentran los archivos "priori\_x\_grids" y "priori\_y\_grids" que permiten la construcción de las funciones de densidad de probabilidad a priori, lo que quiere decir que funcionan como entradas de información previa de los modelos a entrenar en los scripts anteriormente mencionados.

predicción "Estimacion" e "Imputacion".

-  **Resultados:** esta carpeta cuenta con dos subcarpetas
  - **Imputación:** Que son los resultados de la construcción del modelo del script "Proyecto\_de\_grado\_imputacion" y funcionan como entradas de el script "Evaluacion\_imputacion"
  - **Imputación por estimación:** Que son los resultados de la construcción del modelo del script "Proyecto\_de\_grado\_estimacion" y funcionan como entradas de el script "Evaluacion\_estimacion".
-  **Transformaciones:** Por último se encuentra la carpeta que contiene el archivo de las transformaciones relacionadas a cada variable, esta carpeta es útil para los script de predicción "Estimacion" e "Imputacion".