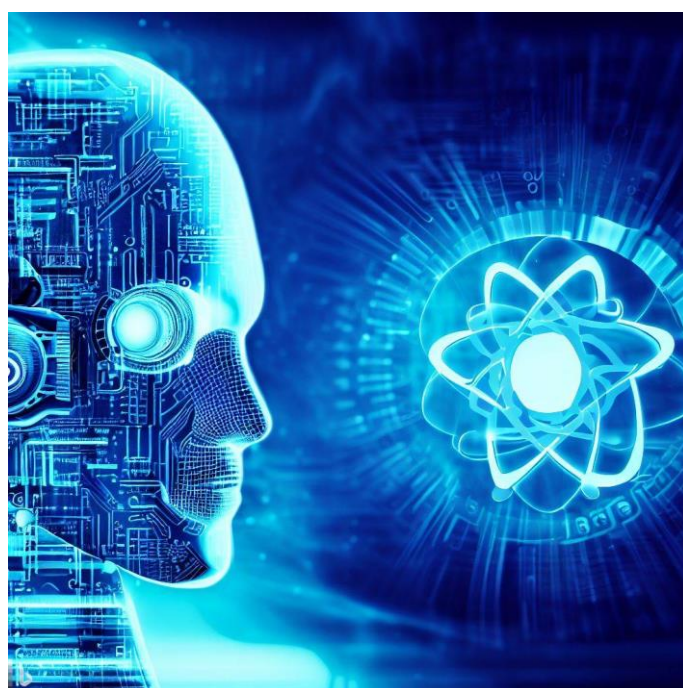


MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

# TRABAJO FIN DE MÁSTER

## ***INVESTIGACIÓN DE MODELOS SUBROGADOS DE QUEMADO DE COMBUSTIBLE BWR BASADOS EN REDES NEURONALES***



**Estudiante:** Suárez Hernández, Ibon

---

**Director:** Hueso Ordóñez, César

---

**Curso:** 2022-2023

**Fecha:** Bilbao, a 29 de mayo de 2023

eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

BILBOKO  
INGENIARITZA  
ESKOLA  
ESCUELA  
DE INGENIERÍA  
DE BILBAO

## RESUMEN

El aumento del consumo energético en los últimos años ha provocado un avance tecnológico que ha hecho surgir varios problemas con los combustibles fósiles, los cuales son uno de los principales causantes del calentamiento global. Para conseguir reducir las emisiones, se debe buscar una forma de generación de energía limpia. Dentro del grupo de energías con bajas emisiones, la energía nuclear es una opción por la que han optado algunos países, al tratarse de una forma de generación sin emisiones y capaz de producir energía eléctrica de forma continua.

Por otro lado, este avance tecnológico ha traído consigo también la aparición de la inteligencia artificial (IA). Esta tecnología ha llegado a nuestras vidas con el objetivo de facilitar el trabajo al ser humano, y se hace uso de ella en una gran diversidad de situaciones. Dentro del campo de la inteligencia artificial se encuentra el aprendizaje profundo o Deep Learning, método mediante el cual las redes neuronales aprenden a distinguir patrones y características significativas de un problema en concreto, teniendo así la capacidad de acelerar alta variedad de procesos.

En este trabajo se hará uso de esta tecnología con el objetivo de diseñar una red neuronal que acelere el proceso de cálculo de ciertos parámetros de elementos combustibles irradiados para, en un paso posterior no objeto de este trabajo, poder llegar a optimizar el almacenamiento de elementos combustibles en contenedores con la ayuda de un algoritmo genético. Para ser capaces de conseguir lo propuesto, se necesita un set de datos lo suficientemente grande para entrenar de forma correcta la red neuronal y comprobar su precisión.

A fin de conseguir un set de datos de tal tamaño y variedad, se han generado de forma semi-aleatoria características y ciclos de irradiación de elementos combustibles y se ha calculado mediante el programa SCALE, la criticidad y composición de cada uno de ellos tras el enfriamiento, estos valores son los que calculará la red neuronal una vez entrenada. Con el set de datos disponible, se diseña la red, variando el número de capas, las neuronas por capa y otra serie de parámetros referentes al diseño, ajuste y entrenamiento de la red de forma que se minimice el error en las predicciones.

Tras la realización del estudio, se podrá observar el tipo de red escogida y se podrán realizar las predicciones necesarias sobre cualquier elemento combustible que ha sido irradiado en un reactor de agua en ebullición.

Palabras clave: Elemento combustible, criticidad, composición, red neuronal, predicciones

## LABURPENA

Azken urteko energia-kontsumoaren gorakadak aurrerapen teknologiko bat eragin du, eta horrek hainbat arazo sortu ditu erregai fosilekin, horiek baitira berotze globalaren eragile nagusietako bat. Emisioak murrizteko energia garbia sortzeko era bat bilatu behar da. Emisio gutxi aske egiten dituzten energien taldearen barruan, herrialde batzuk aukeratu dute energia nuklearra, emisiorik gabeko sorkuntza modu bat eta energia elektrikoa etengabe ekoizteko gai delako.

Bestalde, aurrerapen teknologiko horrek inteligentzia artifizialaren agerpena ere ekarri du. Teknologia hau pertsonen lana errazteko helburuarekin iritzi da gure bizitzara, eta hainbat egoeretan erabiltzen da. Inteligentzia artifizialaren zelaian, ikaskuntza sakona edo Deep Learning metodoa dago. Metodo horren bidez, sare neuronalek arazo jakin baten patroiak eta ezaugarri garrantzitsuak bereizten ikasten dute, eta, horrela, prozesu-aniztasun handia bizkortzeko gaitasuna dute.

Lan honetan, teknologia hori erabiliko da, elementu erregai irradiatuen parametro jakin batzuk kalkulatzeko prozesua azkartuko duen sare neuronal bat diseinatzeko, ondoren, lan honen helburua ez den urrats batean, elementu erregaiak edukiontzietan biltegitratzea optimizatu ahal izateko algoritmo genetiko baten laguntzarekin. Proposatutakoa lortzeko gai izateko, datu-set handi bat behar da sare neuronala behar bezala entrenatzeko eta zehaztasuna egiaztatzeko.

Tamainu eta barietate horretako datu-set bat lortzeko, elementu erregaien ezaugarriak eta irradiazio-zikloak sortu dira erdi ausaz, eta SCALE programaren bidez, hoztu ondoren horietako bakoitzaren kritikotasuna eta konposizioa kalkulatu dira. Balio horiek dira neurona-sareak behin entrenatuta kalkulatu dituenak. Eskuragarri dagoen datu-setarekin, sarea diseinatzen da, geruza kopurua, geruza bakoitzeko neuronak eta sarearen diseinuari, doikuntzari eta entrenamenduari buruzko beste parametro batzuk aldatuz, iragarpenetan errorea minimizatzen.

Azterketa egin ondoren, aukeratutako sare mota ikusi ahal izango da, eta irakiten ari den ur-erreaktore batean irradiatu den edozein elementu erregairi buruzko beharrezko iragarpenak egin ahal izango dira.

Hitz gakoak: Elementu erregai, kritikotasuna, konposizio, sare neuronala, iragarpenak

## ABSTRACT

Technological development has triggered in an increase of electric consumption. Along with this, some problems related to fossil fuels, which are the one of the main reasons of global warming, have arisen. To be capable of reducing emissions, a clean way of generating electrical energy has to be found. Inside the category of low emission energies, some countries have taken the path of nuclear energy because in addition to having low emissions, it is capable of continuously generating electricity.

On the other hand, this technological development has brought about the rise of artificial intelligence. This technology has come to our lives to make them easier, and people use it in many ways. Inside the field of artificial intelligence, there lays Deep Learning, a method where a neural network learns to distinguish significant characteristics and patterns so that it acquires the ability to accelerate a wide range of processes.

This work will make use of the technology mentioned above with the objective of designing a neural network which can speed up the calculation process of certain parameters of spent fuel assemblies so that, in a future project out of the scope of this work, it could optimize the used nuclear fuel canister loading with the help of a genetic algorithm. In order to be capable of doing this, a set of data big enough to train and test the neural network will be needed.

When it comes to building a wide and varied set of data, the burnup of several fuel assemblies semi-randomly generated has been simulated in SCALE, obtaining the value of the criticality and composition of certain isotopes after an also random cooling time. Those will be the values calculated by the neural network once designed. With a complete set of data, the network is designed varying the number of neurons and layers and training it to achieve the lowest error possible.

Once this work is finished, the type of network chosen can be observed and the network will have the capacity of predicting the values of any spent fuel assembly in a Boiling Water Reactor.

Key words: fuel assembly, criticality, composition, neural network, predictions

# ÍNDICE

RESUMEN.....	3
LABURPENA .....	4
ABSTRACT .....	5
Índice de ilustraciones.....	8
Índice de tablas.....	10
Índice de acrónimos .....	11
1. Introducción .....	12
2. Contexto .....	13
Energía nuclear .....	13
Redes neuronales.....	16
3. Objetivos y alcance.....	18
4. Beneficios .....	19
5. Análisis del estado del arte.....	20
Fisión nuclear.....	21
Ciclo de combustible .....	23
Redes neuronales.....	25
Seleccionar el modelo .....	27
Función de activación.....	27
Compilar el modelo .....	30
Entrenamiento.....	31
Evaluar el modelo.....	32
6. Software necesario.....	34
6.1 Python.....	34
6.2 SCALE .....	34
7. Procedimiento .....	36
7.1 Parámetros a variar .....	36
Geometría.....	36
Enriquecimiento .....	39
Gadolinio promedio.....	40
Temperatura y densidad del moderador .....	40
Tiempos de irradiación .....	40
Potencias de irradiación .....	41
Tiempo de enfriamiento.....	41

7.2 Input base .....	41
7.3 Ejecución de casos y aceleración del proceso .....	45
7.4 Datos obtenidos.....	45
7.5 Desarrollo de la red neuronal.....	46
Preparar los datos.....	46
Definir el modelo .....	47
Compilar el modelo .....	48
Ajustar el modelo .....	48
Evaluar el modelo .....	48
Hacer predicciones .....	49
8. Resultados .....	50
8.1 U-235 .....	57
8.2 Pu-239.....	60
8.3 Am-241 .....	63
8.4 Constante de multiplicación efectiva .....	66
9. Planificación del proyecto .....	69
9.1 Descripción de tareas .....	69
9.2 Diagrama de Gantt .....	70
10. Descargo de gastos.....	71
10.1 Horas internas .....	71
10.2 Amortizaciones .....	71
10.3 Gastos .....	71
10.4 Resumen del descargo de gastos .....	72
11. Conclusiones.....	73
12. Bibliografía.....	74

## Índice de ilustraciones

Ilustración 1: Centrales nucleares en Europa .....	14
Ilustración 2: Centrales nucleares en España y su estado .....	15
Ilustración 3: Primer contenedor almacenado en el ATI de la central de Garoña .....	16
Ilustración 4: Ejemplos de uso de la inteligencia artificial .....	16
Ilustración 5: Gráfica de la energía de enlace por nucleón respecto al número de nucleones .....	20
Ilustración 6: Cronología de las generaciones de reactores nucleares .....	22
Ilustración 7: Elementos de una central nuclear .....	23
Ilustración 8: Distribución probabilística de los productos de fisión .....	24
Ilustración 9: Arquitectura de una red neuronal de alimentación directa.....	26
Ilustración 10: Gráfica de la función sigmoide .....	28
Ilustración 11: Gráfica de la función tangente hiperbólica .....	29
Ilustración 12: Gráfica de la función ReLU .....	29
Ilustración 13: Consecuencias de LR muy grande (izquierda) y muy pequeño (derecha).....	31
Ilustración 14: Algoritmo de entrenamiento de una red neuronal.....	32
Ilustración 15: Algoritmo de predicción-corrección utilizado en TRITON.....	35
Ilustración 16: Primera configuración 8x8.....	37
Ilustración 17: Segunda configuración 8x8.....	38
Ilustración 18: Configuración 9x9.....	38
Ilustración 19: Configuración 10x10.....	39
Ilustración 20: Bloque de composiciones.....	42
Ilustración 21: Bloque de celdas.....	43
Ilustración 22: Bloque de los ciclos de quemado .....	43
Ilustración 23: Definición de una unidad.....	44
Ilustración 24: Unidad global y redondeo de esquinas .....	44
Ilustración 25: Insertar una barra combustible en la unidad global .....	45
Ilustración 26: Error en función del número de épocas en la red global .....	52
Ilustración 27: Predicciones frente a valores reales del U-241 en la red global .....	53
Ilustración 28: Predicciones frente a valores reales del Pu-239 en la red global.....	53
Ilustración 29: Predicciones frente a valores reales del Am-241 en la red global .....	54
Ilustración 30: Predicciones frente a valores reales de la $K_{eff}$ en la red global .....	54
Ilustración 31: Predicciones frente a valores reales del U-235 en la red global con 4000 casos.....	55
Ilustración 32: Predicciones frente a valores reales del Pu-239 en la red global con 4000 casos .....	56
Ilustración 33: Predicciones frente a valores reales del Am-241 en la red global con 4000 casos.....	56
Ilustración 34: Predicciones frente a valores reales de la $K_{eff}$ en la red global con 4000 casos .....	57
Ilustración 35: Gráfica del error frente al número de épocas.....	58
Ilustración 36: Predicciones frente a valores reales del U-235 .....	59
Ilustración 37: Predicciones frente a valores reales del U-235 con el set de datos ampliado.....	60
Ilustración 38: Gráfica del error frente al número de épocas.....	61
Ilustración 39: Predicciones frente a valores reales del Pu-239 .....	62
Ilustración 40: Predicciones frente a valores reales y error del Pu-239 con 4000 casos .....	63
Ilustración 41: Gráfica del error frente al número de épocas.....	64
Ilustración 42: Predicciones frente a valores reales del Am-241 .....	65
Ilustración 43: Predicciones frente a valores reales del Am-241 con 4000 casos.....	66



Ilustración 44: Gráfica del error frente al número de épocas .....	67
Ilustración 45: Predicciones frente a valores reales de la $K_{eff}$ .....	68
Ilustración 46: Predicciones frente a valores reales y error en la $K_{eff}$ con el set de datos ampliado .....	68
Ilustración 47: Diagrama de Gantt del proyecto .....	70

## Índice de tablas

Tabla 1: Características de las distintas geometrías (medidas en cm) .....	39
Tabla 2: Parámetros a variar y su rango .....	41
Tabla 3: Comparación de caso con y sin gadolinio .....	46
Tabla 4: Resultados de la red global .....	50
Tabla 5: Comparación de los errores en los set de datos reducido y ampliado.....	55
Tabla 6: Resultados de la red de U-235 .....	57
Tabla 7: Comparación de los errores en los set de datos reducido y ampliado.....	59
Tabla 8: Resultados de la red de Pu-239 .....	60
Tabla 9: Comparación de los errores en los set de datos reducido y ampliado.....	62
Tabla 10: Resultados de la red de Am-241 .....	63
Tabla 11: Comparación de los errores en los set de datos reducido y ampliado.....	65
Tabla 12: Resultados de la red de la $K_{eff}$ .....	66
Tabla 13: Comparación de los errores en los set de datos reducido y ampliado.....	68
Tabla 14: Resumen horas internas .....	71
Tabla 15: Resumen amortizaciones.....	71
Tabla 16: Resumen gastos .....	71
Tabla 17: Resumen del descargo de gastos.....	72

## Índice de acrónimos

BWR	Reactor de agua en ebullición
PWR	Reactor de agua a presión
TFM	Trabajo de Fin de Máster
UE	Unión Europea
PNIEC	Plan Nacional Integrado de Energía y Clima
ATC	Almacén Temporal Centralizado
ATI	Almacén Temporal Individualizado
CSN	Consejo de Seguridad Nuclear
AGP	Almacenamiento Geológico Profundo
IA	Inteligencia Artificial
API	Interfaz de Programación de Aplicaciones
$K_{eff}$	Constante de multiplicación efectiva
SMR	Pequeños Reactores Modulares
MLP	Multiplayer Perceptron
RNN	Red Neuronal Recurrente
CNN	Red Neuronal Convolucional
SGD	Descenso de Gradiente Estocástico
LR	Tasa de aprendizaje
RNSD	Reactor and Nuclear Systems Division
ReLU	Rectified Linear Unit
$UF_6$	Hexafluoruro de uranio
$UO_2$	Óxido de uranio
$Gd_2O_3$	Óxido de gadolinio
$H_2O$	Agua

## 1. Introducción

El trabajo que se presenta en este documento corresponde a la investigación de modelos subrogados de quemado de combustible de los reactores de agua en ebullición (BWR) basados en redes neuronales. Se trata de realizar una serie de simulaciones del quemado de elementos combustibles tipo BWR variando una serie de parámetros, de manera que posteriormente se obtengan los datos necesarios para poder entrenar de forma correcta una red neuronal. Esta, una vez realizado el proceso de entrenamiento, debería ser capaz de realizar por sí sola los cálculos de los resultados obtenidos debidos a las simulaciones, pero con un ahorro de tiempo importante.

El texto comienza realizando una contextualización del proyecto, indicando la situación actual de la energía nuclear junto con las distintas posturas que se están tomando en relación con esta a la hora de establecer las políticas energéticas de los países y llevar a cabo un proceso de descarbonización, además de la situación de las redes neuronales y el crecimiento que han experimentado este tipo de tecnologías novedosas en los últimos años. Una vez dentro de contexto, se procederá a analizar el alcance y los objetivos de este Trabajo de Fin de Máster (TFM), seguido de los beneficios que este estudio puede ofrecer en distintos ámbitos. Más tarde, se realizará el análisis del estado del arte, donde se expondrá el conjunto de conocimientos acumulado dentro de un área específica hasta la fecha, en este caso referente a la energía nuclear, más concretamente la energía de fisión, y las redes neuronales.

Tras el análisis del estado del arte, se procederá a exponer la metodología seguida durante la realización del presente trabajo. Para poder llevarlo a cabo, se necesitan una serie de parámetros que se irán variando de manera que se pueda obtener un número de casos lo suficientemente amplio como para tener en cuenta tanto la mayor parte de situaciones que se puedan dar durante la operación de un reactor como situaciones que no se suelen dar en el quemado de elementos combustibles. Con todos esos datos, se darán una serie de pasos para automatizar y facilitar su procesado, cuyo único objetivo es reducir el tiempo del proyecto. Una vez procesados los datos, se calculará mediante el código SCALE la constante de multiplicación efectiva junto con las concentraciones de algunos isótopos de interés tras un largo periodo de enfriamiento.

Después de calcular los resultados del quemado, se procederá a desarrollar la red neuronal. La gran amplitud del set de datos tiene como objetivo obtener un número de casos suficiente como para realizar un correcto entrenamiento de la red neuronal, y que esta pueda detectar los patrones y las características significativas que influyen en el problema para que posteriormente pueda realizar predicciones lo más precisas posible. Los datos se dividirán en dos grandes grupos de manera que la primera parte sirva para realizar el entrenamiento mientras que la segunda parte se utiliza para comprobar la precisión de la red.

El estudio realizado en este trabajo contiene también la realización de un diagrama Gantt donde se detalla la planificación de las actividades realizadas junto con su duración, seguido de un descargo de gastos donde se describen estos y las amortizaciones del proyecto.

Finalmente, este TFM termina con una serie de conclusiones extraídas durante la realización del trabajo y tras observar los resultados obtenidos.

## 2. Contexto

### Energía nuclear

El continuo avance tecnológico ha llevado en estos últimos años a un crecimiento exponencial del consumo energético en todo el planeta. Hoy en día, son los combustibles fósiles las fuentes principales de generación de energía, sin embargo, estas fuentes traen consigo grandes emisiones de gases de efecto invernadero en su proceso de combustión, siendo uno de los principales causantes del calentamiento global. Actualmente, existe gran preocupación en la sociedad por el cambio climático y la reciente crisis energética, refiriéndose a estos aspectos varios de los objetivos de desarrollo sostenible de la Agenda 2030. Por ello, los distintos países se encuentran en la búsqueda de un mix energético compuesto por fuentes de energía sostenibles que les permitan conseguir los objetivos y que, en la medida de lo posible, les acerque a la independencia energética para que cuestiones ajenas a cada uno tengan la mínima influencia.

Para la consecución de estos objetivos, muchos están optando por potenciar las fuentes de energía renovables, puesto que hay tantas formas, alguna de ellas es accesible para cada país. Sin embargo, se trata de formas de energía no continuas, por lo que deben ser apoyadas por otras formas de energía para abastecer las necesidades de la población en cada momento. Otro camino que están tomando algunos países es el de potenciar la energía nuclear, puesto que se trata de una forma de energía sin emisiones de dióxido de carbono en operación, es continua y garantiza el abastecimiento eléctrico. La energía nuclear es una forma de energía que últimamente está en boca de todos. Se trata de un tipo de energía que tiene sus defensores y sus detractores, viendo que son varios los países que confían en esta forma de generación de energía para la producción de energía eléctrica, mientras que otros, por su parte, se están dedicando a eliminar su rastro cerrando todas las centrales nucleares en operación. Muchos de sus detractores apelan que la problemática de la energía nuclear es que un accidente puede suponer consecuencias catastróficas, y es este miedo a un posible accidente junto con el desconocimiento por parte de la mayoría respecto a la seguridad en las centrales que puede ser la causa principal por la cual se han decidido cerrar las centrales en determinados países.

Hoy en día, en el mundo existen alrededor de 450 reactores en operación divididos en 33 países, siendo Estados Unidos, China y Francia aquellos con un mayor número, produciendo en torno al 10,5% de la electricidad mundial. A diciembre de 2021 había otras 58 unidades en construcción en 20 países, entre los que, junto con China que es el que más construye, se encuentran Rusia, Turquía y Emiratos Árabes Unidos, este último se convirtió en el primer país árabe con energía nuclear, tras el inicio de la operación de uno de sus reactores (Foro nuclear, 2021).

A parte de los reactores que se encuentran en construcción, hay que tener en cuenta que la vida de un reactor está entre 40 y 60 años. Sin embargo, con la presentación de estudios que reflejan la viabilidad de seguir operando de forma segura una vez cumplido su plazo de diseño se puede llevar a cabo la continuidad de operación de las centrales. En Estados Unidos, por ejemplo, seis reactores han recibido la autorización para alargar su vida y operar hasta los 80 años (Foro nuclear, 2021).

Dentro de los reactores, la gran mayoría son de agua ligera, ya sean de agua presurizada (PWR) o de agua en ebullición, y también los hay de agua pesada como los CANDU. A partir del año 2010 se han empezado a diseñar pequeños reactores modulares, los cuales suponen una gran ventaja en zonas alejadas puesto que pueden instalarse en zonas remotas sin conexión a la red, abasteciendo las necesidades energéticas de comunidades alejadas por sí solos (Liou, 2021). Otra ventaja de estos reactores modulares es que la mayoría se están diseñando con sistemas de seguridad pasivos, esto quiere decir que no se necesita la

actuación del ser humano en caso de fallo. En la actualidad se están llevando a cabo diseños innovadores como los refrigerados por plomo o aquellos que utilizan sales fundidas como combustible.

Hablando más concretamente del panorama europeo, la invasión de Ucrania por parte de Rusia ha hecho evidente la necesidad de una mayor independencia energética, junto con esto, los distintos países se encuentran también ante el desafío de la descarbonización, por lo que surge la necesidad de tener un mix energético con las mínimas emisiones posibles, pero que sea capaz de abastecer las necesidades energéticas.

Dentro de la Unión Europea (UE), son 14 de los 27 países miembros los que tienen centrales nucleares, estando 181 en operación y 12 en construcción, siendo el porcentaje de energía consumida de origen nuclear del 25%. Dentro de la UE son Francia, con un 78% de la energía generado a partir de energía nuclear, Eslovenia, con un 54%, y Ucrania, con otro 54%, los países que más confían en esta forma de energía. Se pueden ver las distintas centrales nucleares en operación en Europa (Foro nuclear, 2021).

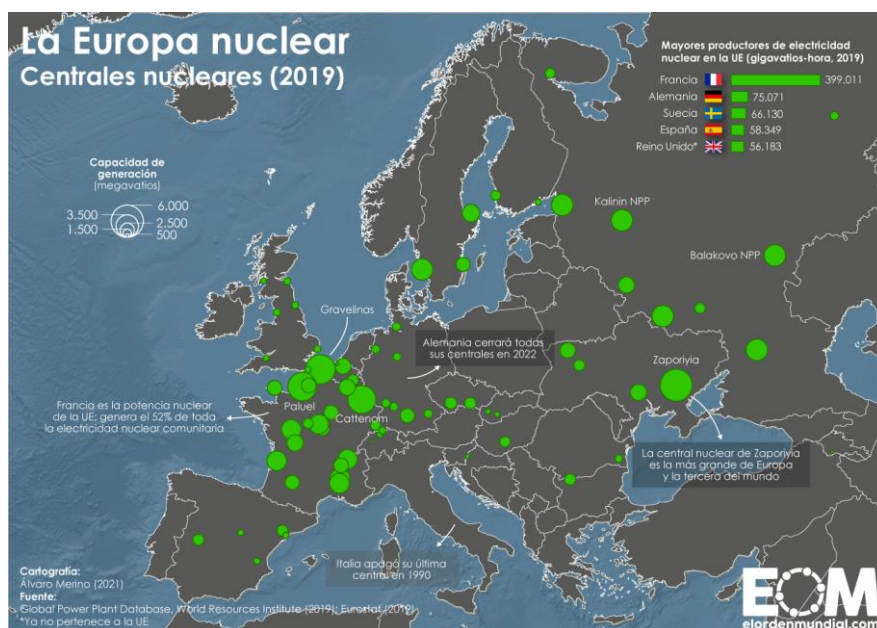


Ilustración 1: Centrales nucleares en Europa

Así como varios países confían en la energía nuclear, son otros los que están en el polo opuesto. Países como Italia, debido al accidente de Chernóbil, y Lituania, por dudas respecto a la seguridad de las centrales, dejaron de operar sus centrales. Además de estos, son otros los que están planeando el cierre. Alemania y Bélgica anunciaron que apagarían todas sus centrales, pero debido a la reciente crisis sus planes se han tenido que retrasar. En la misma línea se encuentran Países Bajos y España, las cuales planean el cierre de sus centrales para el año 2035 (Merino, 2021).

En el caso de España, son 7 el número de centrales que están activas a día de hoy, mientras que la central de José Cabrera se encuentra en proceso de desmantelamiento, la central de Santa María de Garoña está en cese definitivo y Vandellós I se encuentra en periodo de latencia (Ministerio para la transición ecológica y el reto demográfico, 2023), se puede ver en la ilustración 2 su localización y estado. Las centrales activas en España producen en torno al 20% de toda la electricidad consumida y, hasta 2021, fue la principal forma de generar energía sin emisiones, siendo recientemente superada por la eólica. La energía de origen nuclear es esencial en la estabilidad, equilibrio y regulación del sistema eléctrico, además de ser esencial en la transición hacia una economía descarbonizada.

A pesar de que las nucleares pueden ayudar a cumplir con los objetivos propuestos por el Plan Nacional Integrado de Energía y Clima (PNIEC) de descarbonizar la economía para 2050 y reducir la factura de la electricidad de los ciudadanos, el gobierno, aún teniendo en cuenta que la mayoría de los países se están planteando alargar la vida de sus centrales nucleares en vista de la crisis energética, ha decidido mantener el plan inicial de cierre, donde se irán cerrando todas las centrales entre 2027 y 2035 (Pinar, 2022). Las principales razones por las cuales el gobierno ha decidido reducir la dependencia de la energía nuclear son la preocupación por la seguridad y la gestión de residuos radiactivos, así como el miedo a un posible accidente. En consecuencia, con esta política y viendo el desarrollo de las tecnologías renovables, se ha decidido invertir los recursos en este tipo de fuentes de energía.

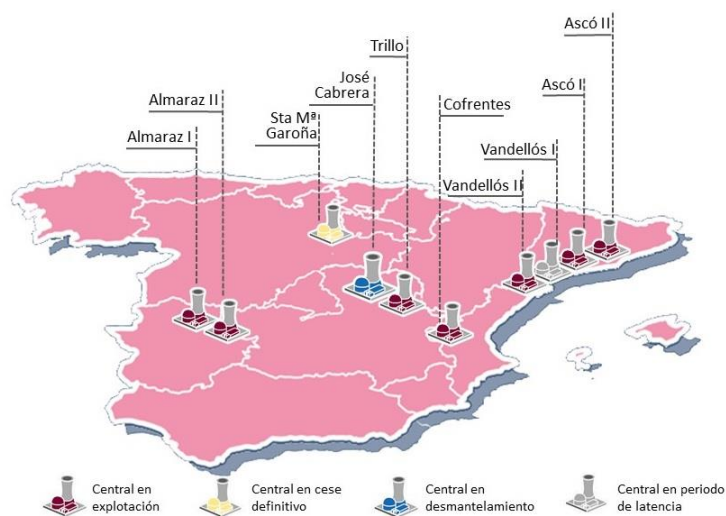


Ilustración 2: Centrales nucleares en España y su estado

Como se ha mencionado previamente, la gestión de los residuos es una de las mayores preocupaciones de aquellos que son contrarios a la energía nuclear, en especial con los residuos de alta actividad y el combustible gastado.

Los elementos combustibles, una vez irradiados, se quedan en las piscinas de combustible de la propia central durante un periodo mínimo de 5 años. Una vez salen de las piscinas, pasan a almacenarse en seco en contenedores que se pueden llevar a un Almacén Temporal Centralizado (ATC) o a un Almacén Temporal Individualizado (ATI), los cuales son instalaciones diseñadas para el almacenamiento temporal de residuos de alta y media actividad. En el caso de los primeros, se almacenan residuos provenientes de varias centrales, mientras que los segundos se ubican junto a la propia central y almacenan los residuos provenientes de ella.

En España, en un principio, se decidió en 2011 construir un ATC en la localidad conuense de Villar de Cañas. Sin embargo, el proyecto ha recibido a lo largo de los años distintas demandas judiciales y ha sido objeto de controversia, por lo que finalmente se decidió pararlo. Finalmente, según el último borrador del Plan General de Residuos, y avalados por el Consejo de Seguridad Nuclear (CSN), se ha optado por que cada central disponga de su propio ATI y, posteriormente, se construirá un Almacén Geológico Profundo (AGP) donde almacenar de forma definitiva los residuos, este, según las predicciones, estará operativo para 2073. En el caso de España en particular, se realiza un ciclo abierto de combustible, por lo que no se lleva a una planta de reprocesamiento para aprovechar el elemento combustible. En la ilustración 3 se puede ver un ejemplo de ATI, el de la central de Garoña (EDCM, 2022).

Cuando se trata de almacenar los elementos combustibles en contenedores, además de ser necesario proporcionar un blindaje frente a la radiación, se debe garantizar la subcriticidad de estos a lo largo del tiempo. Es en este punto donde este TFM presenta especial interés, puesto que, para evaluar la criticidad de elementos de combustible gastado tras el enfriamiento, se hacen necesarias simulaciones complejas que son computacionalmente caras y requieren mucho tiempo.



Ilustración 3: Primer contenedor almacenado en el ATI de la central de Garoña

## Redes neuronales

La inteligencia artificial es algo que estos últimos años ha experimentado un crecimiento brutal, hasta el punto en que, hoy en día, son utilizadas por una gran variedad de empresas y organizaciones en campos muy diversos, desde la tecnología y salud hasta las finanzas y el comercio. En la ilustración 4 se pueden observar distintos usos que se le puede dar a la inteligencia artificial (Observatorio de Inteligencia Artificial, 2016).



Ilustración 4: Ejemplos de uso de la inteligencia artificial

Seguramente, la aplicación principal de la inteligencia artificial y redes neuronales que todo el mundo conoce es el asistente personal que utilizan empresas como Apple y Amazon mediante los productos Siri y Alexa respectivamente. Estos sistemas utilizan herramientas de reconocimiento de voz y procesamiento del lenguaje natural de forma que pueden resolver las peticiones de los usuarios. Es tal la implicación de Apple en esta tecnología que a lo largo de los últimos años ha ido creando una serie de startups especializadas



como puede ser Perceptio, que pretende mejorar la seguridad de los dispositivos y la privacidad de datos mediante la ejecución de sistemas de inteligencia artificial en el propio teléfono, sin hacer uso de la nube. Otro ejemplo es Faceshift, la cual capta las expresiones faciales de las personas y crea animaciones que imitan dichas expresiones.

Otra de las empresas con mayor impacto en el sector de la inteligencia artificial es Google, utilizando esta tecnología en la mayoría de sus herramientas. El buscador de Google utiliza la IA para ofrecer una búsqueda personalizada mediante las sugerencias o el autocompletado. En YouTube, también se hace uso de ella para las recomendaciones de videos, o en Gmail para los filtros de spam, entre otros. Gracias a que esta empresa ha comenzado a liberar sus interfaces de programación de aplicaciones (API), los distintos desarrolladores pueden utilizar de forma sencilla dichas herramientas. Dentro de todo esto, se encuentra Tensorflow, una librería de código libre para aprendizaje automático o Machine Learning desarrollada por Google, de forma que permite construir y entrenar redes neuronales artificiales, y es la herramienta de la que se hará uso en este trabajo.

A parte de estas y otras de las empresas del sector tecnológico, las redes neuronales también se emplean en sectores como el de la salud, para poder analizar grandes conjuntos de datos y establecer una serie de patrones que permitan detectar anomalías y desarrollar tratamientos más efectivos. En el sector de los negocios también es interesante su uso para predecir fluctuaciones en los mercados y generar modelos de riesgo más precisos. Pero no son solo las empresas las que hacen uso de la IA, los desarrolladores e incluso los entusiastas de esta tecnología también hacen uso de ella con fines de investigación o aplicaciones prácticas.

Estos últimos años, el campo nuclear está investigando el uso de la IA para generar modelos subrogados que sustituyan cálculos complejos que llevan mucho tiempo, entre otros, un estudio el cual trata de optimizar la carga en bidones de combustible gastado utilizando una red neuronal y un algoritmo genético (Solans, y otros, 2021).

### 3. Objetivos y alcance

El trabajo objeto de este informe consta de dos fases diferenciadas, además de otra fase que se puede realizar en paralelo con las otras y corresponde a la investigación y recopilación de información para mejorar en la realización del proyecto. En la primera fase, se procederá a realizar las simulaciones necesarias del quemado de elementos combustibles mediante códigos de cálculo del paquete SCALE, variando una serie de parámetros. Específicamente, se utilizará TRITON, que es un módulo de control desarrollado en el marco de SCALE. Esta fase trata de generar un set de datos lo suficientemente grande y variado, que comprenda la mayoría de las situaciones en las que puede operar un reactor.

En la segunda fase del proyecto, una vez ejecutados todos los casos y extraída la información necesaria de los archivos de salida, se procederá a diseñar una red neuronal, la cual será entrenada con el objetivo de que pueda predecir los valores de los parámetros de interés. Para que la red funcione de forma correcta, las predicciones se deben acercar lo máximo posible a los valores que se obtendrían mediante la simulación, solo que el proceso de cálculo es mucho menor.

La primera fase culmina con el cálculo de los parámetros de interés, para el caso que se plantea en este trabajo, se limitará a calcular la constante de multiplicación efectiva ( $K_{eff}$ ), y las concentraciones de algunos isótopos importantes para la criticidad, más concretamente, el Uranio-235 (U-235), el Plutonio-239 (Pu-239) y el Americio-241 (Am-241). Durante la segunda fase del proyecto, en el diseño de la red neuronal, el objetivo será tratar de reducir el error en la predicción lo máximo posible, tanto el error de entrenamiento como el de prueba de la red. Para conseguirlo, en la primera fase se ha debido generar un set de datos lo suficientemente grande.

El presente proyecto solamente se centrará en calcular los parámetros mencionados en elementos combustibles de reactores del tipo BWR, con geometrías 8x8, 9x9 y 10x10, similares a los elementos combustibles de las centrales de Cofrentes, todavía en activo, y Santa María de Garoña, la cual está en situación de cese definitivo.

El proyecto, como se ha dicho previamente, alcanza hasta el cálculo de los valores en cada elemento combustible. Un paso posterior podría valerse de este trabajo y poder calcular los valores de la constante de multiplicación efectiva de un contenedor de almacenamiento y, mediante un algoritmo, poder optimizar la carga de contenedores.

## 4. Beneficios

Al hablar de los beneficios que trae consigo la realización de este trabajo, es oportuno comenzar por el más evidente de ellos, que es el ahorro en el tiempo de cálculo de los parámetros de interés. Una vez desarrollada la red neuronal y habiendo verificado que las predicciones realizadas son correctas o que el error de la predicción es asumible, los usuarios posteriores tardarán apenas unos segundos en realizar el cálculo, teniendo en cuenta que, a través de la simulación, la duración está en torno a la hora. Además, se pueden realizar predicciones de varios casos a la vez mientras que, cuando se simula, si no se dispone de varias unidades de cálculo, solamente se pueden realizar de uno en uno.

También, al suponer este proyecto un primer paso hacia la consecución de una red neuronal capaz de optimizar la carga de contenedores, otros usuarios que se adentren en esta tarea podrían hacer uso de la red neuronal diseñada en este trabajo para, junto con un algoritmo de optimización, conseguir optimizar la carga de contenedores. De esta manera, al mejorar dicho proceso, el uso de recursos y materiales de almacenamiento podrá ser menor y más seguro.

Otro beneficio que puede tener este trabajo es la profundización en el conocimiento de las redes neuronales. Últimamente, su uso ha crecido de forma exponencial, por lo que una buena comprensión de los elementos utilizados en este trabajo para diseñar la red neuronal como son Python y Tensorflow, permitirá en un futuro el diseño de redes neuronales con otros fines. Esto puede ser muy beneficioso en un futuro puesto que no es común encontrar ingenieros con un conocimiento transversal en dos campos, IA y nuclear.

## 5. Análisis del estado del arte

La energía nuclear es la energía contenida en el núcleo de un átomo, los cuales son las partículas más pequeñas en las que se puede dividir un elemento químico manteniendo sus propiedades. El núcleo de los átomos está formado por dos clases de partículas, los protones y los neutrones, que se mantienen unidas debido a la energía nuclear fuerte. Los protones tienen carga positiva, y su número es característico de cada elemento químico. Por otro lado, los neutrones son partículas sin carga, dentro de un mismo elemento puede haber átomos con distinto número de neutrones, los cuales se denominan isótopos. Todos los isótopos de un elemento tienen las mismas propiedades químicas, pero las propiedades físicas pueden diferir de uno a otro. Una de las propiedades físicas que difiere entre los isótopos es la reactividad, pudiendo existir la posibilidad de haber isótopos que emiten radiación, a los que se conoce como inestables, y otros que no, los llamados estables (Consejo de Seguridad Nuclear, 2022).

Una de las formas más útiles de aprovechar la energía nuclear es la generación de energía eléctrica, para ello, es necesario liberar la energía. En la ilustración inferior se puede ver una gráfica que representa la energía de enlace por nucleón respecto del número de nucleones donde, cuanto mayor sea la diferencia entre el o los átomos iniciales y finales, mayor será la energía liberada. Para conseguir liberar dicha energía, se pueden seguir dos caminos distintos, la fisión y la fusión. Por un lado, la fisión consiste en partir de un núcleo muy pesado, aquellos situados a la derecha de la gráfica, y fraccionarlo para obtener como producto núcleos más ligeros, esta reacción es la que se da en los reactores nucleares. Por otro lado, se encuentra la fusión, donde se parte de núcleos muy ligeros para obtener como producto otro más pesado y energía. Esta reacción es la que se da en las estrellas y que es posible debido a la atracción de la fuerza gravitatoria, sin embargo, en la tierra, todavía se están estudiando formas de poder conseguir esta reacción de manera que se pueda disponer de una fuente prácticamente ilimitada de energía con nulas emisiones en operación y con una importante reducción de la cantidad y peligrosidad de residuos en comparación con la fisión.

Como se puede apreciar en la gráfica, el isótopo del hierro Fe-56 es el de mayor energía de enlace por nucleón, por lo que con núcleos más pesados interesará la fisión mientras que con núcleos más ligeros habría que optar por la fusión para poder conseguir energía.

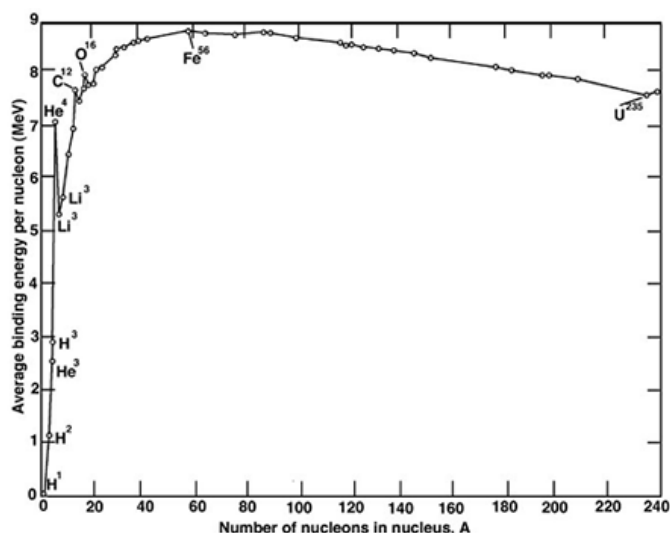


Ilustración 5: Gráfica de la energía de enlace por nucleón respecto al número de nucleones

Tanto en la reacción de fisión como en la de fusión se experimenta una pérdida de masa, la cual se convierte en energía calorífica y de radiación, y se puede conocer su valor gracias a la ecuación de

equivalencia entre masa y energía de Einstein ( $E = mc^2$ ). El fin principal de la tecnología nuclear desde el principio fue para generar energía eléctrica, aunque a lo largo del siglo XX y XXI se ha descubierto que se puede utilizar con otros fines, buenos en el caso del sector de la salud mediante el tratamiento y detección de enfermedades, y malos en los casos de armas de destrucción masiva.

## Fisión nuclear

Como se ha dicho previamente, la fisión nuclear consiste en la división de un núcleo pesado al capturar un neutrón incidente, resultando en dos elementos más ligeros y la emisión de otros productos, como pueden ser los neutrones. Como resultado del proceso, también se emite radiación gamma y grandes cantidades de energía.

Al capturar el núcleo un neutrón, este se vuelve inestable y, para tratar de lograr una situación de estabilidad, se divide en fragmentos más ligeros, y así, ascender en la curva que se muestra en la ilustración 5. Además de los fragmentos ligeros, también se emiten nuevos neutrones, proceso que se conoce como multiplicación, estos ayudarán a producir nuevas reacciones, repitiendo el proceso, obteniendo así una reacción en cadena, la cual es lo que se busca en las centrales nucleares. Dentro de las reacciones de fisión, la  $K_{eff}$  indica la relación entre el número de neutrones en una generación y la siguiente, y mide la eficiencia de la reacción. Si es mayor que uno, se produce más de un neutrón, lo que hace que la reacción se multiplique de forma exponencial, lo que se considera una situación supercrítica, si es menor que 1 la reacción se irá extinguiendo, subcrítico, y si es igual a 1 la reacción es crítica, donde la tasa de producción y la tasa de destrucción de neutrones es la misma.

Las centrales nucleares son el lugar donde se realiza la producción de energía eléctrica a partir de reacciones de fisión en cadena. Estas no son más que una central térmica, donde se calienta un fluido para, posteriormente, generar vapor y expandirlo en la turbina. La diferencia es que el estado térmico del fluido se produce a través del calor térmico que se genera en las reacciones de fisión, es decir, la diferencia radica en el foco calorífico, y es esta la que implica que las centrales nucleares y las térmicas convencionales tengan puntos de vista de diseño y operación distintos, exigiendo los primeros mecanismos de control, barreras de confinamiento y sistemas de seguridad y emergencia para prevenir o mitigar accidentes (Hueso, 2022).

Fue en los años 50 cuando se comenzó a utilizar esta forma de generación de energía eléctrica, y, a lo largo de los años se ha ido evolucionando en el diseño de los reactores. Todo comienza con los primeros prototipos, llamados los reactores de Generación I. Es a partir de los años 70 cuando se empieza a generar energía de forma comercial mediante los reactores de Generación II, que en su mayoría corresponden los reactores de agua ligera, los BWR y PWR, y los CANDU de agua pesada. La vida útil para estos reactores estaba en torno a los 60 años, por lo que muchos de estos están llegando a su final, sin embargo, se están llevando a cabo diversos proyectos para la ampliación de la vida de estos.

Los reactores de Generación II de los que se ha hablado corresponden a la mayoría de los que hoy en día están activos, seguidos de los de Generación III, que comenzaron a operar en la década de los 90 y son fruto de la continua investigación en el campo de la energía nuclear. Dentro de este grupo aparecen reactores avanzados de agua ligera y pesada. Siguiendo con el proceso de investigación y mejora continua aparece la Generación III+. Es en esta etapa donde, además de mejorar el diseño de los reactores que ya estaban en operación, surge una nueva tendencia, los pequeños reactores modulares (SMR), los cuales la mayoría de ellos se basan en sistemas de seguridad pasivos, donde no sea necesaria la actuación de un ser humano para, en caso de accidente, poder solventarlo de la forma más segura posible. Otro gran punto a

favor de estos es que se pueden utilizar en zonas alejadas y de difícil acceso para dar abastecimiento a poblaciones sin tener que estar conectados al sistema eléctrico. La tendencia de la investigación en este campo sigue su curso y hace ver que en el siguiente paso se llevarán a cabo diseños innovadores de reactores, los de Generación IV, aunque el futuro puede estar ligado a los SMR debido a las ventajas que pueden aportar (Foro Nuclear, 2021).

En la siguiente imagen se pueden ver los resultados de la investigación continua mediante una cronología donde aparecen los distintos tipos de centrales desde los inicios, junto con varios ejemplos de centrales en cada generación.

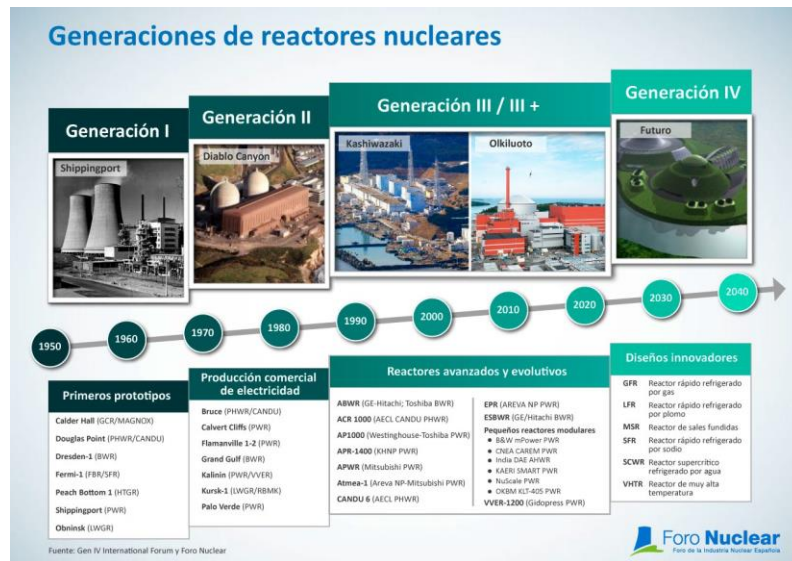


Ilustración 6: Cronología de las generaciones de reactores nucleares

A pesar de que haya distintos tipos de reactores, en la gran mayoría de ellos los elementos principales son los mismos. En una central nuclear, el principal componente es el reactor, donde se producen las fisiones, se libera el calor y es recogido por el refrigerante. El reactor está alojado en una vasija de acero, y el núcleo está formado por elementos combustibles. El combustible alojado en los elementos es lo que va a producir las fisiones, este, además de como fuente de energía, también se utiliza como fuente de neutrones para poder mantener la reacción en cadena de forma controlada. El combustible va alojado en vainas de manera que los productos de fisión se mantengan confinados y no amenacen la seguridad de la central. Para ejercer control sobre el núcleo se utilizan barras de control, normalmente fabricadas con una aleación de Indio, Cadmio y Plata o de carburo de boro, materiales con gran capacidad de absorción neutrónica ya que, controlando los neutrones, se controlan también las fisiones. Dentro de la vasija, en los reactores PWR antiguos también se puede encontrar el presionador, que mantiene en equilibrio la fase líquida y la fase vapor para controlar la presión del reactor, en los actuales ya no se encuentra dentro de la vasija sino en el circuito primario (Foro Nuclear, 2022).

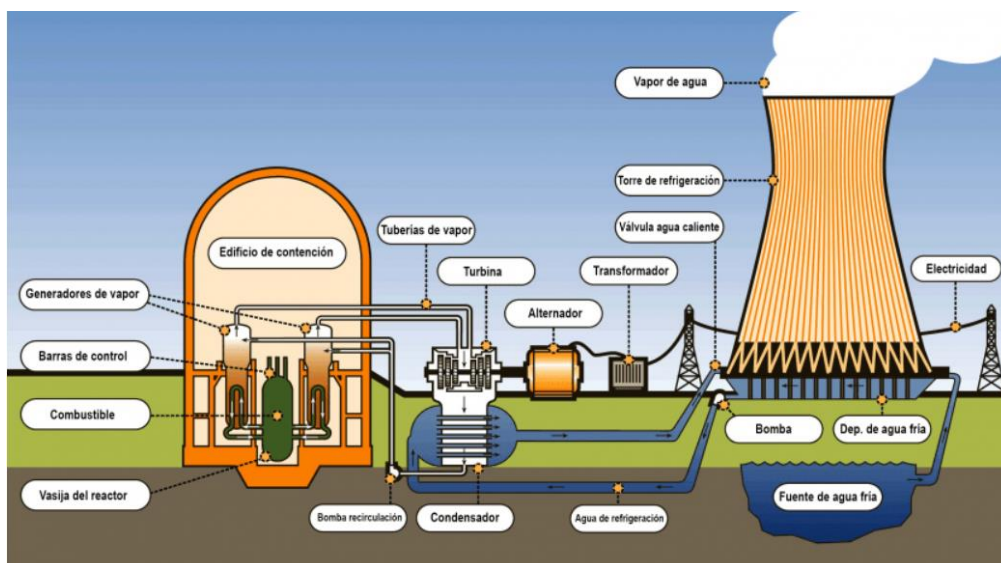
Dentro del reactor, para que se produzcan las fisiones en este tipo de centrales, puesto que la fisión es mayor a energías bajas, será más fácil producir fisiones si los neutrones son térmicos, pero esto no quiere decir que no sea posible mantener la reacción en cadena con neutrones rápidos, como en los reactores rápidos. Al producirse una reacción de fisión, los neutrones se producen con mucha energía, por lo que se necesita algo que los frene, denominado moderador. El moderador debe ser de un material ligero de manera que los neutrones experimenten choques elásticos y vayan gradualmente perdiendo energía. Por el

reactor también circula el refrigerante, el encargado de absorber el calor disipado en las fisiones y generar vapor en el caso de las centrales PWR o ir directamente a la turbina en las BWR. En estos tipos de reactores, el agua refrigerante también actúa como moderador debido al hidrógeno que encontramos en la molécula de H<sub>2</sub>O.

Fuera del reactor se pueden encontrar los generadores de vapor en el caso de las centrales PWR, ya que en las centrales BWR el vapor se genera en el propio núcleo y va directo a la turbina. Cuando el vapor llega a la turbina, se expande, obteniendo así energía mecánica de rotación, la cual se convierte en energía eléctrica en un alternador cuyo eje va solidariamente unido al eje de la turbina. Como las condiciones en las que se genera la energía eléctrica no son óptimas para el transporte, se necesita un transformador para elevar la tensión y que de esta manera las pérdidas en el transporte sean lo más pequeñas posibles.

Una vez sale el vapor de la turbina, pasa al condensador donde vuelve a su estado líquido al intercambiar calor con el agua proveniente de un río, un embalse o incluso del mar.

Por último, destacan los sistemas de seguridad y control, utilizados para garantizar el correcto funcionamiento de la central. Además, no hay que olvidarse del edificio de contención, el cual es una estructura formada por gruesos muros de hormigón armado que alberga la vasija con el reactor o el sistema de refrigeración entre otros. En la siguiente ilustración se pueden apreciar las distintas partes de una central tipo PWR, lo que las diferencia de las centrales BWR son los generadores de vapor.



*Ilustración 7: Elementos de una central nuclear*

### Ciclo de combustible

Anteriormente se ha dicho que es el combustible el elemento principal para generar la energía mediante las fisiones. En la gran mayoría de casos, en los reactores de agua ligera BWR y PWR, el combustible está formado por óxido de uranio, enriquecido en uranio-235, y se encuentra en estado sólido en forma de pastillas cerámicas. En reactores más modernos se utilizan sales fundidas, mezclas de óxidos de uranio y plutonio o las llamadas partículas TRISO (Office of Nuclear Energy, 2019). También se puede utilizar uranio natural como combustible en reactores de agua pesada, esto es porque el agua pesada es mejor moderador de neutrones que el agua ligera y hace que los neutrones producidos se absorban en el

combustible de forma más eficiente, y como la captura de neutrones por parte del combustible es mayor, se necesitará menor cantidad U-235.

El uranio en estado natural se encuentra con un 0,71% de U-235, el cual es fisiónable. Para poder alcanzar la cantidad necesaria se debe enriquecer en este isótopo. Durante el proceso de enriquecimiento, el mineral de uranio se va purificando en distintas disoluciones y precipitaciones hasta conseguir hexafluoruro de uranio ( $UF_6$ ). De ahí, se somete al proceso de enriquecimiento. A pesar de existir varios métodos para poder llevar este proceso a cabo, el más extendido es el de la centrifugación, donde, pasando por distintas etapas, a la salida se obtiene un óxido de uranio enriquecido. Los valores nominales tradicionales para los reactores PWR y BWR están entre un 2,5 y un 5% de enriquecimiento, puesto que una cantidad menor a la mínima no alcanza para producir el número necesario de fisiones mientras que, por su parte, una cantidad superior al límite no se suele utilizar por motivos de licencia. A pesar de todo, hoy en día se fabrica combustible nuclear con enriquecimientos de hasta el 20% para reactores modernos, y es a partir de ese valor que se pueden generar sospechas de estar fabricando combustible con fines armamentísticos.

Tras el paso del combustible por el reactor, no se puede tratar inmediatamente como residuo debido a su capacidad de emitir calor residual y la presencia de isótopos radiactivos. El uranio, al fisiónar, da lugar a productos de fisión, y debido a la naturaleza estocástica de la reacción de fisión, cada reacción puede dar lugar a distintos productos de fisión, sin embargo, en la ilustración 8 se puede apreciar una distribución probabilística de ellos. Pero no son solo los productos de fisión los que interesan, sino que también hay que tener en cuenta los actínidos, producidos por la absorción neutrónica en isótopos pesados ya sean el Uranio (como el Pu-239 producido por la absorción en U-238) o productos de activación como el Np-237 que se produce por doble captura (de U-235 pasa a U-236 y este a Np-237) o por reacción ( $n, 2n$ ) del U-238, que se caracterizan por su larga vida.

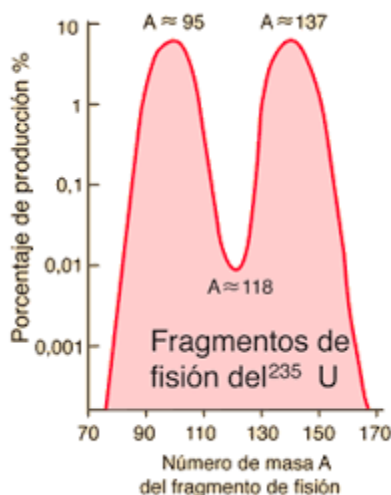


Ilustración 8: Distribución probabilística de los productos de fisión

Con el objetivo de poder llevar a cabo una gestión del combustible gastado de forma correcta, de manera que no suponga una amenaza contra las personas o el medio ambiente, se debe seguir un proceso. Según un elemento combustible sale del reactor tras ser irradiado por última vez, se inserta en las piscinas de combustible, donde deberá permanecer un periodo mínimo de enfriamiento que suele oscilar entre 3 y 5 años. Las piscinas se encuentran en las propias centrales, de manera que no haya que trasladar el combustible en ese estado, y tienen como función principal refrigerar los elementos combustibles extrayendo el calor residual, debido principalmente a las desintegraciones gamma de los productos de



fisión de vida corta. El agua, además, supone un buen blindaje biológico contra la radiación emitida y, unido al hecho de ser transparente, lo hacen idóneo para permitir operaciones de inspección y realizar un manejo seguro de los elementos por parte de los trabajadores.

Tras un tiempo en las piscinas, cuando el combustible pierde gran parte de su capacidad de generar calor residual y su radiactividad, se puede considerar como residuo, aunque todavía hay que tenerlo controlado. La siguiente etapa de la gestión del combustible gastado transcurre en los almacenes temporales, donde se almacenan los elementos combustibles en seco. Estos, como su nombre indica, tienen como función almacenar el combustible y otros residuos de alta actividad durante un tiempo determinado para, posteriormente, llevarlos a una instalación de reproceso en el caso de los que sigan un ciclo cerrado, o a su etapa final de almacenamiento en un AGP. Las instalaciones de almacenamiento temporal suelen estar situadas en la superficie o a poca profundidad, pudiendo almacenar combustibles de una sola central en el caso de los ATI, o de varias centrales si se trata de un ATC. En ambos casos se requieren medidas de vigilancia continuadas, siendo el ATC la solución más idónea en términos de seguridad puesto que las medidas de vigilancia y control están más centralizadas.

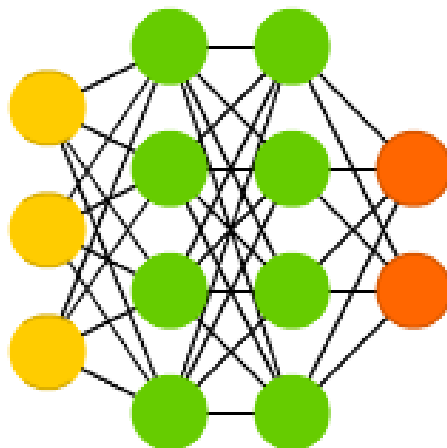
Una vez transcurrido el tiempo en el almacén temporal, queda la última etapa en el proceso de gestión de combustible gastado. Para esta etapa, la solución considerada más favorable por la comunidad hoy en día es la del AGP, el almacenamiento a grandes profundidades en instalaciones situadas en formaciones geológicas estables. La solución consiste en interponer una serie de barreras en las instalaciones situadas entre 500 y 1000 metros bajo tierra, con los residuos insertados primero en contenedores de metal extremadamente resistentes a la corrosión y otras formas de degradación, después, se sitúan en unas galerías taponadas con materiales absorbentes y en un terreno con una muy alta capacidad de retención y muy poco permeable, de manera que se aisle de la mejor manera posible y no suponga ningún peligro para las generaciones venideras ni para la naturaleza. En términos de seguridad física, esta opción es también la más adecuada al ser poco vulnerable al robo y sabotaje.

Para este tipo de almacenamiento. Existen distintas formaciones geológicas donde se pueden realizar, siendo las mejor estudiadas las arcillas, la sal y las rocas duras magmáticas o metamórficas, entre las que destacan el granito, el gneis, el basalto o la toba. Sobre todo, se buscan zonas estables y que las rocas no tengan fallas o estén poco fracturadas.

La característica principal que define esta etapa es que ya no se pretende recuperar el material residual y no es necesaria la realización de controles activos de la instalación puesto que esta se cierra y se sella. Sin embargo, se puede mantener vigilada tanto la instalación de almacenamiento como el medio ambiente a su alrededor (Consejo de Seguridad Nuclear, 2019).

## Redes neuronales

Las redes neuronales son un tipo de modelo de aprendizaje automático inspirado en lo que se conoce del comportamiento del cerebro humano, no se trata de imitarlo porque resulta tarea imposible. Estas consisten en un conjunto de nodos interconectados, llamados neuronas, que procesan y transmiten la información. Para poder realizar su función de forma correcta, las redes neuronales se organizan en capas, en The Assimov Institute se pueden observar las distintas arquitecturas de las redes neuronales, aunque una de las más utilizadas, y la empleada en el desarrollo de este proyecto, se muestra en la ilustración 9, consistente en una red profunda de alimentación directa.



*Ilustración 9: Arquitectura de una red neuronal de alimentación directa*

Dentro de las capas se distinguen 3 tipos, la capa de entrada, que debe tener el mismo número de neuronas que los datos de entrada, la capa de salida y las ocultas.

Las capas de las redes neuronales están formadas por neuronas, las cuales son la unidad funcional de las redes. Dentro de cada neurona, se produce la suma ponderada de todas las entradas y la aplicación de la función de activación. Al realizar la suma ponderada de las entradas, se multiplica cada uno de los valores de entrada por su peso y se realiza la suma, a esto, se le debe sumar también el sesgo de cada neurona. Después, al valor obtenido se le aplica la función de activación obteniendo así el valor de salida de la neurona. La función de activación sirve sobre todo para no obtener salidas que sean simplemente combinaciones lineales y poder obtener una gran diversidad de salidas.

El uso que se le da a las redes neuronales es el de automatizar tareas intelectuales normalmente llevadas a cabo por seres humanos, sobre todo en resolución de problemas como clasificación de imágenes, reconocimiento de voz o predicción de series temporales entre otros. La capacidad de las redes neuronales para aprender patrones complejos y no lineales las hace interesantes para estos cálculos ya que los sistemas de ecuaciones para el quemado es un sistema no lineal.

A fin de obtener buenos resultados, las redes utilizan un proceso de entrenamiento supervisado donde se les proporciona un conjunto de datos de entrada y un conjunto de etiquetas de salida correspondientes. Durante el entrenamiento, la red ajusta sus parámetros (los pesos y los sesgos) con el objetivo de minimizar el error entre las etiquetas de salida y los valores de salida reales. Una vez que se ha completado el entrenamiento, la red puede utilizarse para hacer predicciones sobre nuevos datos de entrada.

Para diseñar de forma correcta una red neuronal, es necesario seguir el denominado ciclo de vida de una red. En primer lugar, se debe definir el modelo junto con la arquitectura, la cual consiste en el número de capas y las neuronas por capa. Posteriormente, se compila el modelo, en este paso se debe seleccionar una función de coste a minimizar y un algoritmo de optimización, pudiendo también establecer una serie de métricas además de la función de coste para evaluar las predicciones. Una vez compilado el modelo, el siguiente paso es ajustarlo, es decir, entrenarlo, y para ello, se deben seleccionar los parámetros de entrenamiento como las épocas o el tamaño de lote. Tras el ajuste del modelo, se procede a su evaluación, para ello, se utilizan datos que el modelo no ha visto antes. Por último, tras realizar todos los pasos anteriores, el modelo está listo para ser utilizado y realizar las predicciones necesarias (Brownlee, 2020).

A continuación, se detallarán un poco más los distintos pasos del ciclo de vida de las redes neuronales.

### Seleccionar el modelo

En el momento de seleccionar un modelo, lo más común es utilizar uno secuencial, donde se define una clase secuencial y se pueden ir añadiendo capas de una en una, conectando todas las capas de la misma forma. También existen los modelos funcionales, que se tratan de modelos más avanzados donde se da la posibilidad de conectar la salida de una capa con la entrada de otra con una conexión específica.

Dentro de un mismo modelo, es posible seleccionar distintos tipos de redes. Las Multiplayer Perceptrons (MLP) están compuestas por una o más capas de neuronas, alimentando los datos a la capa de entrada, mientras que en la o las capas ocultas se realizan las correspondientes operaciones para realizar las predicciones en la capa de salida. Este tipo de redes son adecuadas para problemas de regresión y clasificación cuando a las entradas se les asigna una clase siendo también muy recomendable su uso para problemas en los que los datos estén dispuestos en forma de tabla.

Las Redes Neuronales Convolucionales (CNN) están diseñadas para cualquier tipo de predicción cuya entrada sea una imagen. Además, aunque no hayan sido específicamente diseñadas para ello, también han arrojado buenos resultados en la clasificación de documentos.

Por último, están también las Redes Neuronales Recurrentes (RNN), que han sido diseñadas para trabajar con problemas en los que interviene el factor tiempo, es decir, en predicciones de secuencias. Se incluyen secuencias de texto y de lenguaje hablado, campo en el que tienen mucho éxito, sobre todo en lo que se conoce como procesado natural del lenguaje. Este tipo de redes son muy difíciles de entrenar, y no son aconsejables para datos dispuestos en forma de tabla o para imágenes como datos de entrada.

Los tipos de CNN y RNN no se suelen utilizar por sí solos, lo más común es utilizar capas correspondientes a estos modelos junto con una o más capas del modelo MLP, generando modelos mixtos. Estos tipos pueden también apilarse en determinadas arquitecturas como modelos de reconocimiento (Negm, 2019).

### Función de activación

A las neuronas se les debe aplicar una función de activación de forma que las salidas no sean simplemente combinaciones lineales. Dentro de las funciones de activación, se deberá escoger aquella que sea adecuada para el problema en cuestión. La función de activación es una función matemática que se le añade a cada neurona para que la red pueda aprender patrones complejos en los datos. Comparándola con el cerebro humano, la función de activación decide qué se va a disparar a la siguiente neurona. Junto con todo lo anterior, también ayudan a mantener el valor de la salida en un cierto límite para que no se den lugar problemas de cálculo.

Hay que decir que no cualquiera puede actuar como función de activación. Para que se pueda utilizar con este fin, debe cumplir la fuga de gradiente, es decir, que cuando se aplique el descenso del gradiente y la propagación hacia atrás, que el gradiente de cada capa afecte al anterior. Por otra parte, la función debe ser simétrica para que los gradientes no se desplacen en una dirección. También debe tener un gasto computacional bajo ya que al aplicarse después de cada capa se deben calcular un gran número de veces. Por último, tienen que ser diferenciables, puesto que es necesario poder derivarlas para poder aplicar el método del descenso del gradiente durante el entrenamiento (Ponce, 2021).

Dentro de las funciones de activación, las que forman parte del grupo de las funciones de activación lineal no son utilizadas en aprendizaje profundo porque la salida no está limitada. Las de activación no lineal son las más utilizadas, dentro de las cuales destacan las siguientes:

- **Función sigmoide:** Como se puede apreciar en la imagen, tiene forma de S y está centrada en 0,5 con un rango en el eje de ordenadas entre 0 y 1. Normalmente, se suele aplicar en capas finales de un problema de clasificación binaria. Se trata de una función diferenciable, pero, aunque esta sea monótona, su derivada no lo es.

El problema de esta función es que, al tener entradas pequeñas o grandes, se centra en 0 o 1, por lo que la derivada será cercana a 0, y entonces, el gradiente no se propaga por la red. A esto se le suele añadir un gasto computacional alto al tener que realizar operaciones exponenciales.

$$f(x) = \frac{1}{1 + e^x} \quad (Ec 1)$$

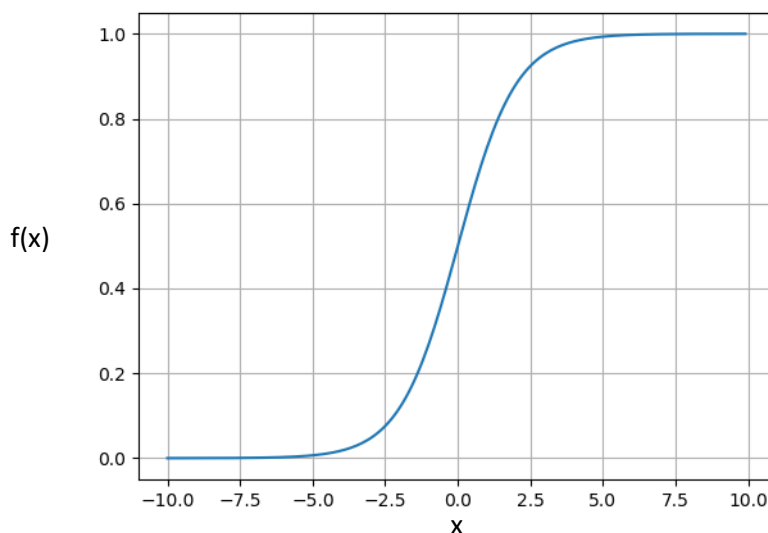


Ilustración 10: Gráfica de la función sigmoide

- **Función tangente hiperbólica:** En la ilustración 11 se puede ver que es bastante similar a la función sigmoide, con la salvedad de que en este caso la función se centra en el 0, con un rango en el eje de ordenadas entre -1 y 1.

Para esta función, se dan los mismos problemas que en el caso anterior.

$$f(x) = \tanh(x) \quad (Ec 2)$$

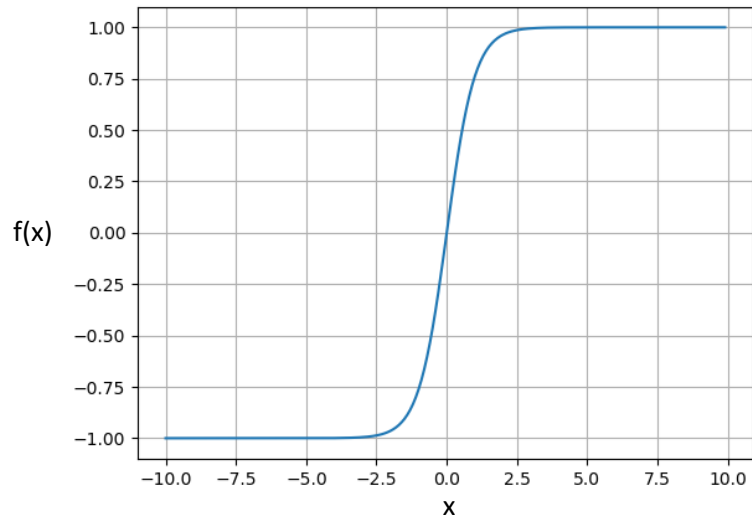


Ilustración 11: Gráfica de la función tangente hiperbólica

- Función ReLU: Es la función más utilizada debido a que tiene una mejor propagación del gradiente, un cálculo eficiente al estar hablando de operaciones simples y supone un coste temporal bajo con una escala invariante.

$$ReLU(x) = \max(0, x) \quad (Ec 3)$$

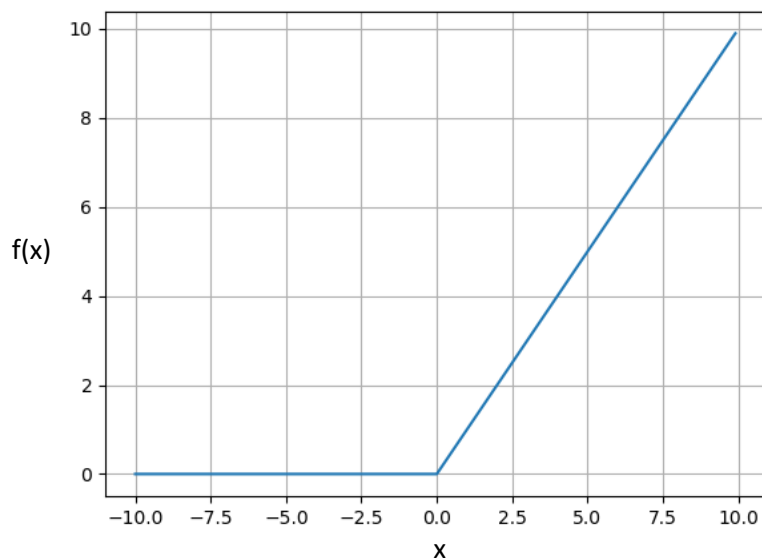


Ilustración 12: Gráfica de la función ReLU

Entre los problemas que acarrea esta función se encuentran que no está centrada en 0. Además, en 0 la función no es diferenciable, pudiendo escoger como valor de la derivada el 0 o el 1. También se puede dar la muerte de algunas neuronas cuando la suma ponderada de los pesos da como resultado un valor negativo.

Para resolver algunos de estos problemas, han surgido algunas funciones derivadas de esta, como la función PReLU, donde se inserta una pequeña pendiente en la zona negativa de las abscisas para asegurar que la neurona nunca muera. Se ha demostrado que esta función trabaja mucho mejor en problemas con muchos datos.

- Función Softmax: Se trata de una forma de la función sigmoide utilizada en problemas de clasificación de múltiples clases. Los parámetros de salida estarán en el rango entre 0 y 1, siendo la suma de todos ellos la unidad. También se utiliza como capa final en problemas de clasificación, pero en este caso no solamente de clasificación binaria.

## Compilar el modelo

A la hora de compilar el modelo, se necesita establecer una función de coste o error, donde el objetivo de la red va a ser minimizarla en el entrenamiento para que, una vez desarrollada, al realizar las predicciones, se acerque lo máximo posible a la realidad. Junto con la función de coste, es necesario seleccionar un algoritmo de optimización. En un principio se utilizaba el descenso del gradiente, pero a medida que aumentaba el número de capas y neuronas comenzaba a generar problemas. Esto sucede porque el hecho de que el gradiente sea cercano a 0 en muchas regiones puede hacer que la optimización se estanque, por lo que, para evitar este problema, se han realizado modificaciones en el algoritmo de manera que se adapte la tasa de aprendizaje. Así, la velocidad del proceso se adapta a las características de cada región.

Dentro de los algoritmos de optimización, el más simple es el de descenso de gradiente estocástico (SGD). Se introduce un comportamiento estocástico debido a la inviabilidad del cálculo de la derivada parcial de la función de coste respecto a cada uno de los pesos. Este algoritmo limita el cálculo de la derivada a solamente una observación por lote. A partir de este han surgido otros algoritmos como el Adaptive Gradient Algorithm (AdaGrad) y su variación Adadelta, los cuales no consideran un valor uniforme para cada uno de los pesos. En su lugar, se mantiene un factor de entrenamiento para cada peso y, como resultaría inviable calcularlo de forma específica, lo va escalando y adaptando. El Root Mean Square Propagation (RMSProp) es otro algoritmo similar, manteniendo un factor de entrenamiento diferente. Por último, se encuentra el algoritmo Adam, Adaptive moment estimation, el cual combina las propiedades de AdaGrad y RMSProp. Los distintos algoritmos de optimización se construyen en base a sus predecesores, por lo tanto, se puede esperar un mejor rendimiento de este último (Velasco, 2020).

La tasa de aprendizaje o Learning Rate (LR) establece la velocidad de aprendizaje del modelo, es decir, cómo de rápido pueden cambiar los parámetros mientras se optimiza. Este parámetro es muy difícil de establecer puesto que depende de los datos y se relaciona con el resto de los parámetros. Un LR grande implica que el modelo puede ir saltando de una región a otra sin aprender, mientras que un LR pequeño alarga el proceso de entrenamiento y puede estancarse, esto se puede observar en la ilustración 13, donde a la derecha se muestra lo que puede ocurrir con un LR pequeño y a la izquierda lo que ocurre con un LR demasiado grande. Lo aconsejable es utilizar un LR pequeño siempre y cuando el tiempo de entrenamiento no supere ciertas limitaciones y no utilizar un LR constante durante el proceso.



Ilustración 13: Consecuencias de LR muy grande (izquierda) y muy pequeño (derecha)

## Entrenamiento

El entrenamiento de la red sirve para poder realizar predicciones realistas en un futuro y consiste en modificar los distintos pesos de las conexiones de las capas de las neuronas.

Para entrenar la red, existen ciertos parámetros que tienen gran importancia en el proceso. Uno de ellos es el tamaño de lote, es decir, el número de ejemplos que se introducen para que entrene cada vez. Si el número es muy pequeño, la red ocupa poca memoria de datos y el entrenamiento es más rápido, sin embargo, existe la posibilidad de que no aprenda del todo ciertas características significativas. Por otro lado, si este es muy grande, sí que tendrá en cuenta todas las características significativas, pero la velocidad del proceso se verá muy afectada.

Otro parámetro muy influyente en el entrenamiento es el número de épocas. Este, se refiere al número de veces que va a pasar cada ejemplo del entrenamiento por la red neuronal. Se puede pensar que un alto número de épocas sería lo ideal, haciendo que la red aprenda, pero una red más compleja no tiene por qué ser mejor, lo mismo sucede con el número de capas y neuronas. En principio, un alto número de estos parámetros puede implicar un mejor aprendizaje, no obstante, esto no es siempre así. El peso y el sesgo de cada neurona van a converger a un valor, y una vez se llega, se va a consumir tiempo y energía sin aprovecharlo, pudiendo también ocurrir el fenómeno del sobreajuste. Este fenómeno se da cuando crece el error en la prueba mientras que el error en el entrenamiento sigue bajando. Lo que ocurre en este caso es que, en vez de aprender patrones y características significativas, la red memoriza los casos que se le han alimentado y, al introducir casos que no ha visto antes, su funcionamiento será peor (González, 2020).

Con el objetivo de realizar el proceso de entrenamiento las redes utilizan los métodos de descenso del gradiente y propagación hacia atrás. La propagación hacia atrás es uno de los bloques más importantes al construir una red neuronal, sirve para entrenar la red de forma efectiva, sabiendo que la red es un apilamiento de neuronas en forma de capas, donde intervienen las funciones de activación de estas, con la intención de reducir el error a la salida.

El algoritmo de entrenamiento se basa en que, después del paso hacia delante de los datos en la red neuronal, este efectúa una propagación hacia atrás mientras se ajustan los parámetros del modelo. La modificación de los parámetros depende de la derivada, y avances en los últimos años han llevado a mejores algoritmos de optimización basados en el gradiente. En la ilustración inferior se puede ver de forma esquemática el proceso de entrenamiento de la red.

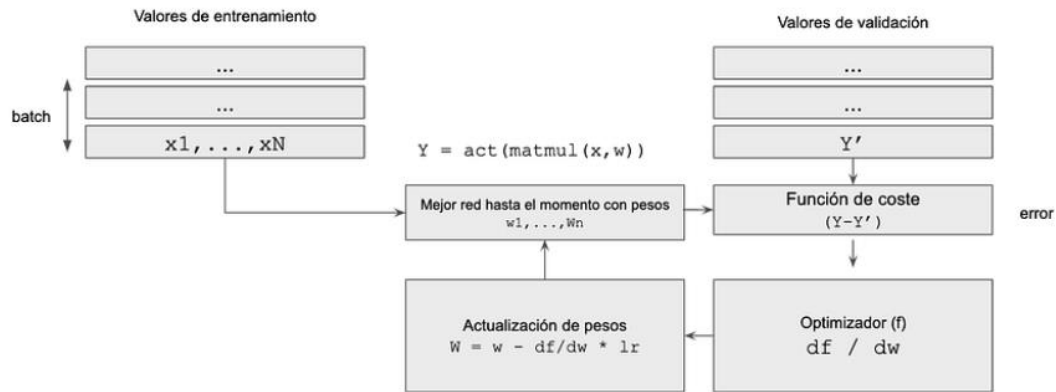


Ilustración 14: Algoritmo de entrenamiento de una red neuronal

Cuanto más parámetros se tengan para ajustar, mayor tendrá que ser el set de datos de entrenamiento. En otras palabras, en la propagación hacia atrás se busca minimizar la función de coste ajustando los pesos y sesgos de la red. El nivel de ajuste está determinado por gradientes de la función de coste con respecto a esos parámetros.

Junto con la propagación hacia atrás, se utiliza también el descenso del gradiente. Esto se hace porque la derivada de una función mide la sensibilidad al cambio del valor de la función respecto a un cambio en el argumento. El descenso del gradiente es un algoritmo de optimización iterativo de primer orden que se utiliza para encontrar un mínimo o un máximo de una función dada, en el caso de Machine Learning se trata de minimizar la función de coste.

El descenso del gradiente no es aplicable a todas las funciones, estas tienen que ser diferenciables, la derivada debe existir en todo el dominio y tienen que ser convexas.

Para entender lo que es el descenso del gradiente, es necesario tener en cuenta qué es un gradiente, la primera derivada de la función en un punto, por eso se dice que es un método de primer orden. En caso de tratarse de una función multivariante, es el vector de las derivadas en cada una de las direcciones principales. Como lo que interesa es una pendiente a lo largo de un eje, no interesan el resto de sus derivadas parciales.

Tras la aplicación de ambos métodos, en el caso de tener derivadas cercanas a cero, la modificación será muy pequeña, por lo que el entrenamiento precisará de mucho tiempo y, por el contrario, con un gradiente elevado, la modificación será mayor y los pesos cambiarán de forma más rápida. Es por esto que, al principio, cuando se establecen los parámetros al azar, los gradientes serán elevados y se ajustará más rápido, pero luego, cuando el valor de los parámetros se esté ajustando a los óptimos, la derivada se acercará a cero y tardará más tiempo en alcanzar el valor deseado.

### Evaluar el modelo

Evaluar el modelo consiste en realizar predicciones con datos que la red no ha visto antes, es por esto que cuando se tratan los datos, se suele reservar una parte de estos para la evaluación. En este paso, no solamente se pasan al modelo las características, sino que también se especifican las etiquetas, las cuales son los resultados que debería dar la simulación.



Para que la evaluación sea visible, existe la posibilidad de establecer una serie de métricas de forma que se pueda comparar el error cometido durante el entrenamiento del error cometido en las pruebas. Se puede observar de esta forma si se ha dado el problema del sobreajuste y cómo evoluciona el error en función del número de épocas.

## 6. Software necesario

Para la realización del proyecto es necesario disponer de ciertos programas. Algunos de ellos servirán simplemente para facilitar o agilizar el trabajo mientras que otros son necesarios tanto para la simulación del quemado de todos los casos como para el diseño de las redes neuronales.

### 6.1 Python

Python es un lenguaje de programación de alto nivel que se utiliza para gran variedad de aplicaciones, desde desarrollo de software hasta aprendizaje automático. El hecho de utilizar Python a lo largo de este proyecto ha facilitado el aprendizaje debido a su sintaxis clara e intuitiva, además del gran número de bibliotecas y herramientas de las que dispone.

A lo largo de la realización de este trabajo, se ha hecho uso de Python en diversas situaciones. En la primera parte, se ha hecho uso de este para automatizar el proceso cuando se trata de cambiar los parámetros en los archivos de entrada, ejecutar los casos y extraer la información necesaria de los archivos de salida, en otras palabras, simplemente para acelerar el proceso. Sin embargo, en la segunda parte del proceso, se ha utilizado para desarrollar la red neuronal, lo que lo convierte en algo esencial.

### 6.2 SCALE

El paquete SCALE (Standardized Computer Analyses for Licensing Evaluation) es un sistema de códigos modular que se combinan entre sí en secuencia para realizar todo tipo de cálculos nucleares. Muy utilizado para el modelizado y simulación adecuado para análisis de seguridad nuclear y diseños que están siendo desarrollados, mantenidos, probados y gestionados por la Reactor and Nuclear Systems Division (RNSD). Proporciona una herramienta completa, verificada y validada para la criticidad, física de reactores, blindaje contra la radiación caracterización de términos fuente radiactivos y análisis de sensibilidad. Dentro del programa, se proporciona un marco integrado con módulos computacionales, incluyendo tres solucionadores de transporte de radiación deterministas y tres de Monte Carlo. También incluye bibliotecas con datos nucleares y herramientas dependientes de los problemas para neutrónica de energía continua y de multigrupos, además de cálculos de activación, agotamiento y desintegración.

Dentro del código SCALE, para la realización del proyecto se utilizará el módulo TRITON, el cual se utiliza para transporte, agotamiento y análisis de sensibilidad e incertidumbre. Además, se puede usar junto con el módulo de agotamiento ORIGEN para predecir concentraciones isotópicas, términos fuente y el calor de decaimiento entre otros.

Dentro de las secuencias que se pueden utilizar, para este problema interesa una secuencia de agotamiento (T-DEPL) donde se automatizan los cálculos de agotamiento o decaimiento tras los cálculos de transporte para cada material especificado. TRITON automatiza los procesos computacionales a través de una serie de intervalos proporcionada por el usuario. Un intervalo de agotamiento representa un intervalo de tiempo en el que la potencia se supone constante, y cada uno puede estar seguido de un cálculo de decaimiento. Dentro de un mismo intervalo, las concentraciones de isótopos en los materiales van a cambiar, por lo que también se experimentarán cambios en la distribución de flujo neutrónico y en las secciones eficaces. Esto lleva al módulo a representar cada intervalo como una serie de intervalos más pequeños.

Para poder procesar el esquema definido por el usuario, TRITON utiliza un método predictor-corrector. Lo que hace este método es realizar cálculos basados en las concentraciones isotópicas anticipadas a mitad de

un subintervalo, después, se realizan los cálculos del subintervalo completo partiendo de lo predicho en el punto medio. Esos cálculos son extendidos al punto medio del siguiente subintervalo, seguido de los cálculos en el nuevo punto medio. Este proceso iterativo se repite hasta que se hayan procesado todos los intervalos (W. A. Wieselquist, 2020).

En la ilustración 15 se muestra un esquema del algoritmo que sigue este módulo, en este caso existen 2 intervalos de agotamiento, el primero de ellos contiene dos subintervalos mientras que el segundo solamente contiene un subintervalo, cada uno de ellos seguido de un intervalo de decaimiento. En la ilustración, las líneas discontinuas significan un cálculo predictor mientras que las otras representan el cálculo corrector.

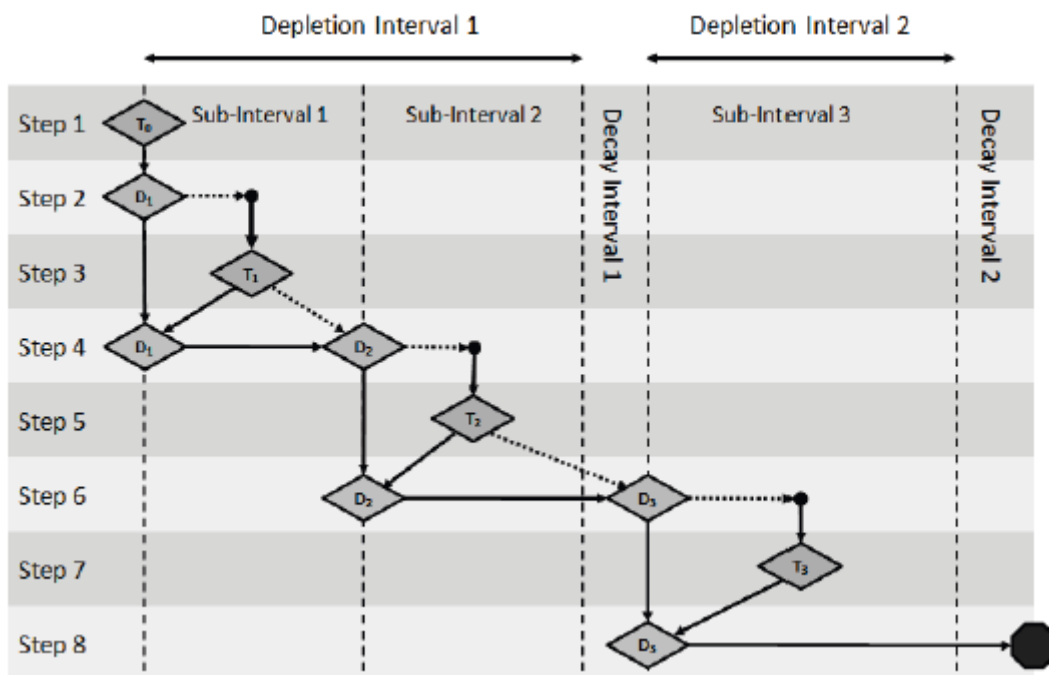


Ilustración 15: Algoritmo de predicción-corrección utilizado en TRITON

## 7. Procedimiento

A lo largo de este apartado, se llevará a cabo una explicación detallada del procedimiento llevado a cabo para la realización de este trabajo, teniendo en cuenta los distintos pasos a dar. Este se dividirá en dos fases, una primera donde se simula el quemado de los elementos combustibles y otra segunda en la que se desarrolla una red neuronal que hará de modelo subrogado.

Una vez obtenido el set de parámetros a variar, se generan un número suficiente de inputs y se ejecutan todas las simulaciones con SCALE para obtener los resultados necesarios. De los archivos de salida, que son muy extensos, solamente interesa para este proyecto el valor de la  $K_{eff}$  y las concentraciones de los isótopos de vida larga de interés: U-235, Pu-239 y Am-241. Con estos datos y los parámetros de entrada, se conformará un set de datos necesario para el entrenamiento de la red neuronal.

En la segunda fase se hará uso del set de datos generado anteriormente para el desarrollo de la red neuronal. Con el objetivo de realizar un entrenamiento correcto, se debe dividir el set en dos, una parte para el entrenamiento y otra para probar la red.

Durante el desarrollo de la red, para minimizar el error, se escogerá la configuración de capas y neuronas adecuada junto con los parámetros de la red necesarios para compilar, entrenar y evaluar la red neuronal. Una vez realizado este proceso, la red neuronal estará lista para hacer las predicciones que serían necesarias para la optimización de la carga de contenedores de combustible gastado.

### 7.1 Parámetros a variar

Con el objetivo de generar distintos inputs de manera que se obtenga un set de datos lo suficientemente amplio como para poder entrenar y testear la red neuronal de forma correcta y que posteriormente realice predicciones adecuadas, será necesario variar una serie de parámetros para tener una gran variedad de casos con los que realizar el entrenamiento.

Para el set de datos, este se generará de forma semi-aleatoria, combinando aleatoriamente los distintos parámetros y filtrando aquellos con un quemado entre 38000 y 50000 MWd/tU, que son valores razonables de quemado de descarga de elementos BWR.

#### Geometría

Lo primero que se va a variar es la geometría. En las centrales nucleares tipo BWR se utilizan diferentes diseños de elementos combustibles con distintas geometrías. En este trabajo se van a tratar 4 diseños distintos con tipos de red: 10x10, 9x9 y dos configuraciones distintas de 8x8. Estas geometrías son las correspondientes a los diseños de la empresa General Electric GE14, GE11, GE4 y GE6, y son algunos de los diseños utilizados en las centrales españolas de Cofrentes y Santa María de Garoña.

Este parámetro es un tanto especial, pues no se va a variar en un mismo archivo de entrada, para este caso, al tratarse de 4 geometrías distintas, se van a establecer 4 plantillas de inputs distintas y a partir de ellos, con la geometría ya definida en cada uno, variar el resto de los parámetros. A continuación, se mostrará en cada caso la geometría y los materiales.

En primer lugar, aparece uno de los elementos 8x8, con un total de 64 barras, correspondiente a la GE4. Dentro del código de colores de la imagen, se pueden distinguir 5 tipos distintos de barras. Dentro de las barras de combustible, las azules, moradas y verdes, tienen la misma composición inicial, están compuestas

por óxido de uranio ( $UO_2$ ) con un determinado enriquecimiento, sin embargo, se deben definir de forma diferente debido a que, debido a su localización en el elemento combustible, el quemado será diferente. Por otro lado, la barra roja corresponde a una barra de agua que ejerce de moderador. Por otro lado, las barras en las que aparecen varios colores formando anillos son las que contienen gadolinio, además de  $UO_2$ . Los venenos consumibles como el gadolinio se utilizan como absorbentes neutrónicos y compensan la reactividad inicial del combustible fresco, permitiendo alargar el ciclo si el quemado del gadolinio se hace de forma correcta. Los anillos que se ven son porque las barras con óxido de gadolinio ( $Gd_2O_3$ ) se modelizan de esta forma debido a que el gadolinio no se gasta de forma homogénea. En los reactores térmicos, puesto que la sección eficaz de absorción es mayor a bajas energías, los neutrones tienen mayor probabilidad de ser absorbidos en las capas externas, por lo que estas blindan a las internas, efecto que se llama auto-blindaje espacial. Este efecto se ve tanto en barras con como sin gadolinio, solo que en las barras con gadolinio es más importante para predecir bien la criticidad y la evolución de los isótopos. El fondo verde corresponde al agua que actúa como moderador y refrigerante que se encuentra en estado bifásico y también se puede distinguir el revestimiento de zircaloy tanto de las barras como del elemento en su conjunto. El revestimiento de las barras de combustible difiere en color de aquel del elemento combustible y de la barra de agua porque el de los primeros es de zircaloy-2 y el de los segundos de zircaloy-4.

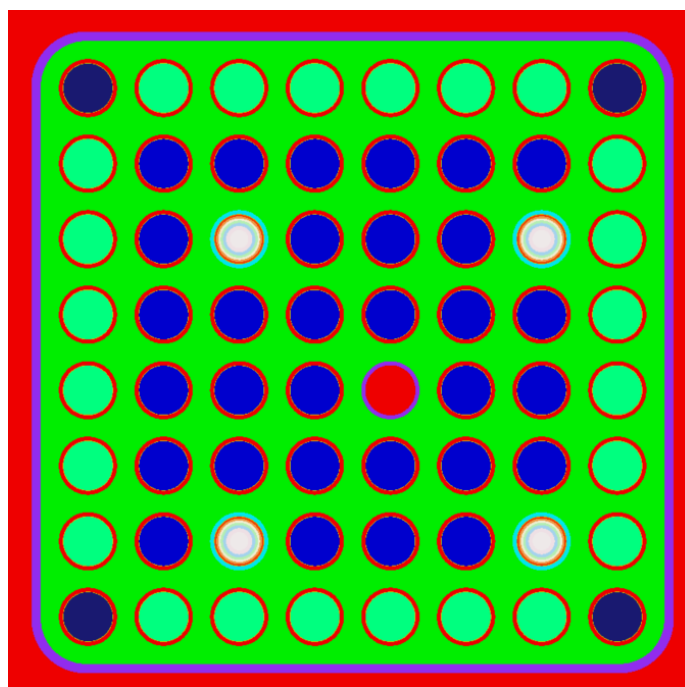
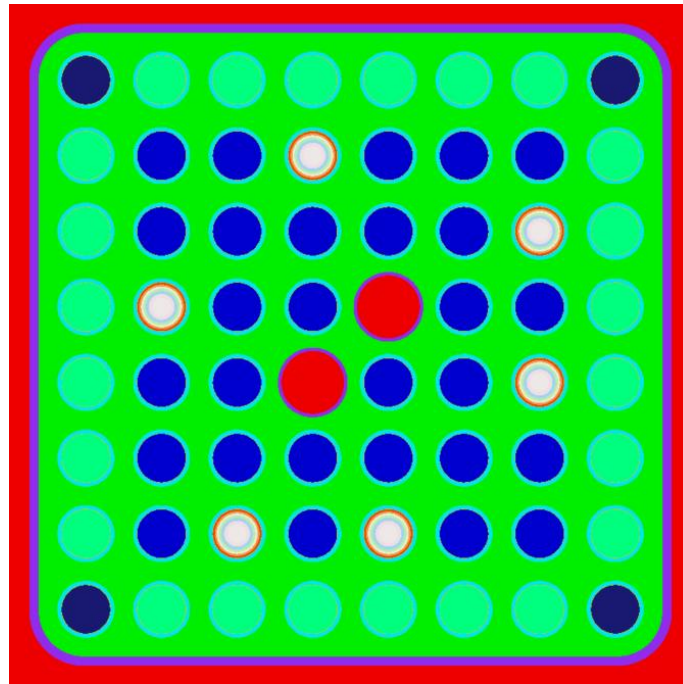


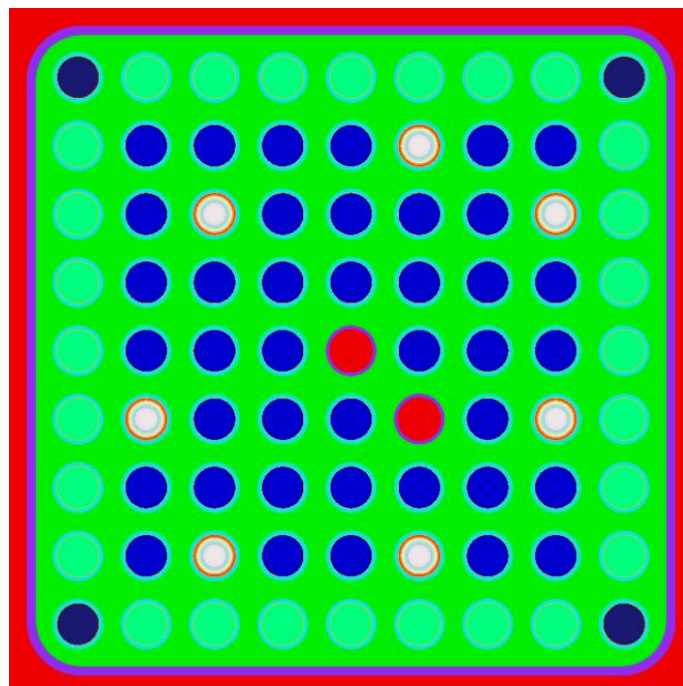
Ilustración 16: Primera configuración 8x8

En segundo lugar, aparece la segunda configuración de 8x8, también con 64 barras, correspondiente a la GE6. En este caso, se diferencia de la configuración anterior debido a que tiene 6 barras con gadolinio y en vez de tener una barra de agua como moderador, aparecen dos. Es fácilmente observable también que el tamaño de las barras de agua es más grande que en la anterior configuración, pero, por otro lado, el tamaño de las barras de combustible es menor, manteniendo el mismo pitch o distancia entre centros y el mismo tamaño de elemento combustible.



*Ilustración 17: Segunda configuración 8x8*

Para el caso del elemento combustible 9x9, que corresponde al elemento GE11 aparecen también dos barras de agua con un tamaño ligeramente mayor que las barras de combustible. El número de barras que contienen gadolinio también es mayor, con un total de 7.



*Ilustración 18: Configuración 9x9*

Por último, se encuentra el elemento combustible 10x10, correspondiente al elemento GE14. Cabe destacar el tamaño de las barras de agua, ocupando 4 posiciones de barras de combustible, y un gran número de barras con contenido de gadolinio de manera que pueda reducir la reactividad inicial que tendrá el combustible fresco.

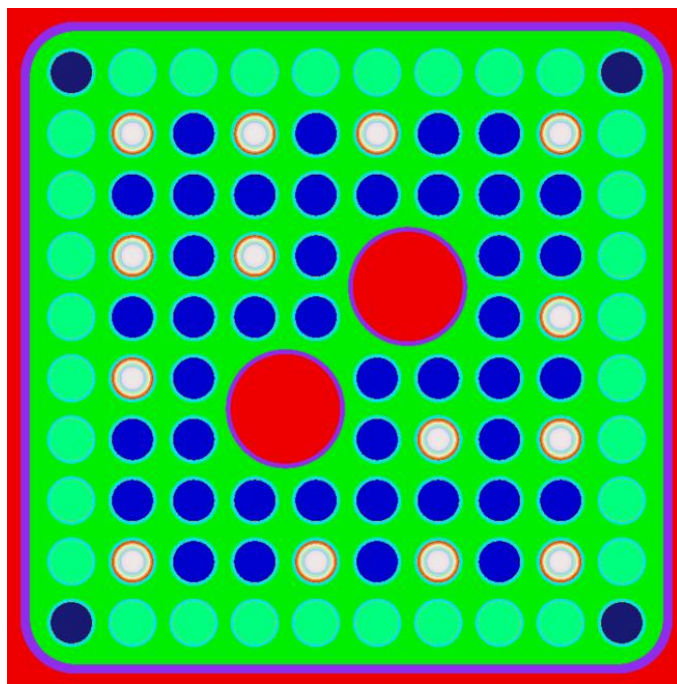


Ilustración 19: Configuración 10x10

En todos los casos, en las barras de combustible tanto con gadolinio como sin él, hay un ligero hueco denominado “gap”, este sirve para permitir la expansión térmica de las pastillas de combustible al ser sometidas a muy altas temperaturas durante la operación y así, no ejercer presión excesiva en el revestimiento, evitando una posible deformación o fractura. Las barras son presurizadas con helio y se deja en la parte superior un espacio denominado “plenum” donde se acomodan los gases que se liberan tras las fisiones. Si se centran las miradas en el elemento combustible se puede observar que hay un eje de simetría, de manera que el proceso de cálculo es más rápido.

En la Tabla 1 se muestra un resumen de las distintas geometrías y sus características.

Tabla 1: Características de las distintas geometrías (medidas en cm)

Geometría	8x8 (1)	8x8 (2)	9x9	10x10
Nº Barras de fuel	59	56	72	78
Nº Barras con gadolinio	4	6	7	14
Diámetro del combustible (cm.)	1,05664	1,0414	0,88456	0,876
Diámetro interno de la barra (cm.)	1,0796	1,06426	0,92456	0,894
Diámetro externo de la barra (cm.)	1,25222	1,22682	1,07696	1,026
Pitch (cm.)	1,6256	1,6256	1,45228	1,2954

### Enriquecimiento

El enriquecimiento se refiere al porcentaje en peso del isótopo de uranio-235 dentro de la cantidad total de uranio. El proceso de enriquecimiento se utiliza para hacer eficaz el combustible nuclear, ya que este isótopo es capaz de producir fisiones a partir de neutrones térmicos. Aunque es cierto que no todos los reactores utilizan uranio enriquecido, los de agua ligera, y entre ellos los BWR objeto de este trabajo sí que lo hacen. Los valores entre los que se va a variar el enriquecimiento serán el 2,5 y el 5%, pudiendo tomar cualquier valor dentro de ese rango.

Es necesario aclarar también que, dentro de una misma barra, la distribución axial del enriquecimiento no es homogénea, siendo esta magnitud en los extremos menor que los valores de las partes centrales. Sin embargo, para no complicar el problema en demasía se tendrá en cuenta que en toda la barra el valor del enriquecimiento será el mismo.

### Gadolinio promedio

Como se ha comentado previamente, el gadolinio actúa como veneno consumible, y sirve para compensar el exceso de la reactividad que tiene el combustible fresco al inicio del ciclo. Para poder ejercer la función que se le atribuye, se necesita una alta sección eficaz de captura de neutrones, la cual poseen los isótopos Gd-155 y Gd-157.

El enriquecimiento en gadolinio de las barras de combustible que lo contienen corresponde al porcentaje en peso del contenido en  $Gd_2O_3$  que hay en la barra, para este proyecto se selecciona entre el 3 o 5%, aunque pueden existir elementos combustibles con un mayor contenido. En este caso se trata de valores discretos y no de un rango.

En la misma línea del caso anterior, esta cantidad no es la misma a lo largo de toda la longitud de la barra, pero para facilitar el modelado del elemento combustible se supondrá constante en la dirección axial.

### Temperatura y densidad del moderador

Otros parámetros que se van a variar son los referentes al moderador, en este caso su temperatura y densidad, estos dos parámetros tienen su influencia en la reactividad y el quemado del combustible.

Las reacciones de fisión se producen en su mayoría debido a la absorción en el combustible de neutrones térmicos, por lo que un cambio en la densidad del moderador va a implicar un cambio en la tasa de absorción de neutrones térmicos. Los neutrones se moderan mediante colisiones elásticas con las moléculas del moderador, por ello se buscan moderadores ligeros. Una densidad baja implica menor número de moléculas de moderador, lo que suponen menores choques y un mayor recorrido libre medio, por lo que el número de neutrones moderados será más pequeño, dando lugar a una menor tasa de absorción de neutrones térmicos. Por otro lado, una alta densidad de moderador implica todo lo contrario, más presencia de moléculas llevan a mayor número de choques, dando lugar a un menor número de fugas al ser el recorrido libre medio menor y más neutrones térmicos.

Para este trabajo, el rango en el que se variará la temperatura será entre 270 y 300°C, mientras que la densidad oscilará entre 0,4 y 0,75 g/cm<sup>3</sup>, siendo en ningún caso valores discretos, sino que pueden tomar cualquier valor dentro del rango. Estos valores corresponden al rango en el que suele operar un reactor BWR.

### Tiempos de irradiación

El tiempo de irradiación es el tiempo que dura un ciclo, es decir el tiempo durante el cual el combustible está siendo quemado dentro del reactor o el tiempo entre recargas parciales. En lo respectivo a este informe, se suponen entre 3 y 4 ciclos, con un periodo de 30 días entre ciclos destinado a la recarga. Este valor se ha escogido por ser el tiempo nominal que suele durar una parada para recargar actualmente.

Para los valores de los tiempos de irradiación, se va a escoger entre tres valores posibles de forma aleatoria. Los ciclos durarán 360, 540 o 720 días. En los casos en los que se ha escogido realizar solamente 3 ciclos, el valor del tiempo de irradiación en el cuarto ciclo sería de 0 días, pero, debido a que en el



programa existe un mínimo para cada parámetro, se ha sustituido por 0,0001 días de manera que se pueda realizar la simulación.

Los 30 días para la recarga comienzan con la parada del reactor y enfriamiento del mismo. Después, se destapa la vasija y se extrae el combustible gastado, se introduce el fresco y se reorganiza el combustible que no estaba en su último ciclo para un mejor aprovechamiento. Este tiempo también se aprovecha para realizar trabajos y pruebas en algunos equipos o sistemas. Finalmente, la recarga termina con el nuevo arranque y sincronización con la red eléctrica.

### Potencias de irradiación

Se trata sobre la potencia específica del elemento combustible para el ciclo correspondiente. Las unidades son MW/tU, o los watos entre las toneladas iniciales de uranio.

El rango de potencias para los distintos casos está entre 0,6 y 5,5 MW por elemento combustible, pero como al programa hay que especificarle la potencia específica, se dividirá el valor correspondiente por 0,175, el cual es la cantidad de uranio inicial.

### Tiempo de enfriamiento

El proyecto trata de calcular el valor de la  $K_{eff}$  y de las concentraciones de ciertos isótopos de interés después de un largo tiempo tras la descarga del combustible. El interés de este parámetro está asociado a asegurar la subcriticidad del combustible durante su almacenamiento, y en una evolución de este trabajo, elaborar un método de optimización de la carga de contenedores que tenga en cuenta la optimización del valor de la  $K_{eff}$  del bulto o contenedor.

Para los tiempos de enfriamiento, se variará en un rango entre 100 y 1000 años, por lo que se calcularán los parámetros de elementos combustibles que estén cerca de salir de las instalaciones de almacenamiento en seco (ATI o ATC) y de los que se trasladen a un almacenamiento definitivo (AGP).

En la Tabla 2 se muestra un resumen de los parámetros que se van a variar y su rango.

*Tabla 2: Parámetros a variar y su rango*

<b>Parámetro</b>	<b>Tipo de variación</b>	<b>Rango</b>	<b>Unidades</b>
Enriquecimiento promedio	rango	2,5-5,0	%
Gadolinio promedio	discreto	3,0-5,0	%
Temperatura moderador	rango	270-300	°C
Densidad moderador promedio	rango	0,4-0,75	g/cm <sup>3</sup>
Tiempos de irradiación	discreto	360-540-720	días
Potencias de irradiación	rango	0,6-5,5	MW
Tiempo de enfriamiento	rango	100-1000	años
Quemado	rango	38000-50000	MWd/tU

## 7.2 Input base

Tal y como se ha dicho en el apartado anterior, los parámetros arriba mencionados se van a variar partiendo de cuatro inputs base, cada uno con su geometría. A continuación, se realizará una explicación de los archivos generados.

En las primeras tres líneas del programa se definen los parámetros y la secuencia del programa, se le asigna un nombre y la biblioteca de secciones eficaces a utilizar. En este caso, la secuencia utilizada será una de agotamiento para conseguir el objetivo del trabajo.

Después, se definen los materiales que van a encontrarse en el problema. Es en este bloque de composiciones donde se van a variar gran parte de los parámetros. Se define la composición de las barras con y sin gadolinio, el enriquecimiento, a qué temperatura va a operar, el material de revestimiento y su temperatura, la densidad y temperatura del agua que actúa como moderador y refrigerante etc. Se puede ver que hay ciertas palabras entre corchetes, las cuales corresponden a los parámetros que se van a variar. Con el objetivo de realizar la variación de los parámetros de forma automática, se desarrollará un programa que sustituya las palabras clave por los valores especificados. En la ilustración 20 se muestra el bloque de composiciones con todo lo definido en este.

```

read comp
' fuel enriquecido
uo2 $fuel den=10.42 1 948.45 92234 0.04 92235 {enriquecimiento} 92238 {u238} end
' combustible con gadolinio
uo2 $fuelgad den=10.42 {porcuo2} 948.45 92234 0.04 92235 {enriquecimiento} 92238 {u238} end
gd2o3 $fuelgad den=10.42 {egad} 948.45 end
' cladding
zirc2 $clad den=6.56 1 636.06 end
' refriger
h2o $refrig den={densmod} 1 {tmod} end
'hueco de helio
he-4 $hueco den=0.00011 1 792.26 end
' moderador
h2o 5 den={densmod} 1 {tmod} end
' tubo de agua
zirc4 6 den=6.56 1 {tmod} end
end comp
  
```

*Ilustración 20: Bloque de composiciones*

El siguiente bloque a definir son las celdas, donde destaca el hecho de que haya tres celdas que son iguales salvo en el número de material. Estas celdas corresponden a las barras de combustible sin gadolinio, pero al localizarse en zonas distintas del elemento combustible, centro, esquinas y laterales, se han tenido que definir tres tipos de celdas distintos, siendo el número de material que se coloca en cada diámetro corresponde al de su alias. La última zona especificada es la de las barras de combustible con contenido en gadolinio, en este caso se define una región en forma de anillos debido a que el gadolinio no se quema de forma homogénea en toda la pastilla, por lo que la evolución del material con el quemado será distinta en función de la distancia al centro. En la ilustración 21 se puede ver que las tres primeras celdas son iguales en cuanto a las medidas, pero es necesario definir las tres al estar en distinto lugar y que el material evoluciona de forma distinta, a pesar de que la composición inicial es la misma.

```

read celldata
  latticecell squarepitch fuel=1.05664 1
                                gapd= 1.0796 4
                                cladd=1.25222 2
                                pitch=1.62560 3 end
  latticecell squarepitch fuel=1.05664 11
                                gapd= 1.0796 14
                                cladd=1.25222 12
                                pitch=1.62560 13 end
  latticecell squarepitch fuel=1.05664 21
                                gapd= 1.0796 24
                                cladd=1.25222 22
                                pitch=1.62560 23 end
  multiregion cylindrical right=white end
    101 0.19969
    111 0.28240
    121 0.34587
    131 0.39937
    141 0.44651
    151 0.48913
    161 0.52832
    34 0.53975
    32 0.62611
    33 0.91715
  end zone
end celldata
  
```

*Ilustración 21: Bloque de celdas*

Posteriormente se define el bloque de agotamiento, donde se seleccionan qué materiales van a variar durante la irradiación. En el problema que se trata a lo largo de este trabajo, no interesa la evolución del revestimiento ni del refrigerante, y es por esa razón que solamente es necesario especificar el combustible, tanto el que contiene gadolinio como el que no.

Siguiendo el orden de los bloques definidos, ahora es el turno de especificar los ciclos a los que se va a ver sometido el combustible dentro del reactor. Tanto la potencia específica como el periodo de irradiación son parámetros a variar, no así el tiempo de parada destinado a la recarga, el cual será de 30 días, que es el tiempo típico de recarga de una central. Con el fin de obtener el valor de los parámetros necesarios tras un largo tiempo de enfriamiento, se le añade un quinto ciclo virtual con un tiempo de irradiación y una potencia bajas para poder realizar los cálculos y el tiempo de enfriamiento se inserta como el tiempo de parada entre los ciclos 4 y 5. Todo esto se puede apreciar en la ilustración 22.

```

read burndata
  power={power1} burn={irrad1} down=30 end
  power={power2} burn={irrad2} down=30 end
  power={power3} burn={irrad3} down=30 end
  power={power4} burn={irrad4} down={enfriamiento} end
  power=0.00001 burn=2 down=1 end
end burndata
  
```

*Ilustración 22: Bloque de los ciclos de quemado*

Seguido, se encuentra el módulo de postprocesado y análisis de resultados, donde se especificarán los distintos isótopos acerca de los cuales se quiere obtener información.

El último gran bloque define el modelo, este se puede dividir en otros más pequeños donde se especifican los parámetros para el desarrollo de la simulación, los materiales que intervienen en el problema, la geometría al completo del elemento combustible y las fronteras. En el caso del problema en cuestión, las fronteras se especifican de forma que todas sean reflectoras, de manera que se simula el quemado de combustible como si estuviera rodeado de otros elementos combustibles iguales, de manera que los neutrones que se escapan por ellas van a ser los mismos que le llegarían por parte de los elementos combustibles que se encuentran alrededor.

Para la geometría se definen unidades distintas que se van a insertar dentro de una unidad global para conseguir el elemento combustible al completo. En la ilustración 23 se muestra cómo se define una barra de combustible tanto con gadolinio (derecha) como sin él (izquierda), especificando la geometría, los materiales que intervienen y las fronteras. Para la barra con contenido en gadolinio, de forma que se consiga la estructura de anillos mencionada, se definen un mayor número de cilindros.

```

unit 1
  cylinder 10 0.52832
  cylinder 20 0.53975
  cylinder 30 0.62611
  media 1 1 10
  media 4 1 20 -10
  media 5 1 30 -20
  boundary 30

unit 4
com='fuel rod - with 3.0% Gd20'
  cylinder 10 0.16717
  cylinder 20 0.23641
  cylinder 30 0.28954
  cylinder 40 0.33433
  cylinder 50 0.37379
  cylinder 60 0.40947
  cylinder 70 0.44228
  cylinder 80 0.46228
  cylinder 90 0.53848
  media 101 1 10
  media 111 1 20 -10
  media 121 1 30 -20
  media 131 1 40 -30
  media 141 1 50 -40
  media 151 1 60 -50
  media 161 1 70 -60
  media 4 1 80 -70
  media 2 1 90 -80
  boundary 90
  
```

Ilustración 23: Definición de una unidad con y sin gadolinio

Para la unidad global, se define el elemento al completo y se redondean las esquinas, de forma que se asemeje más a la realidad, en la ilustración 24 se muestra la definición de la unidad global y el redondeo de una de las esquinas.

```

global unit 10
  cuboid 10 14.32305 0.91695 14.32305 0.91695
  cuboid 20 14.52625 0.71375 14.52625 0.71375
  cuboid 30 15.24 0 15.24 0
  'redondeamos las esquinas
  cylinder 11 0.9652 chord -x=1.88215 chord -y=1.88215 origin x=1.88215 y=1.88215 sides=16
  cylinder 12 1.1684 chord -x=1.88215 chord -y=1.88215 origin x=1.88215 y=1.88215 sides=16
  cuboid 13 1.88215 0.71375 1.88215 0.71375
  
```

Ilustración 24: Unidad global y redondeo de esquinas

Por último, para meter las barras de combustible en la unidad global, se consigue mediante “huecos” como se puede apreciar en la ilustración 25, especificando tanto la unidad a insertar como el origen de coordenadas del hueco.

hole 1 origin x= 6.80720 y= 3.55600

*Ilustración 25: Insertar una barra combustible en la unidad global*

### 7.3 Ejecución de casos y aceleración del proceso

Cuando ya se dispone del set de datos con los parámetros a variar y los inputs base están conformados, mediante el uso del programa Python se escribe un código capaz de generar un archivo input diferente para cada caso sustituyendo convenientemente los parámetros. Una vez se disponga de los distintos archivos, el código los ejecuta en paralelo.

Problemas con el servidor de cálculo han llevado a que las simulaciones se realicen en un ordenador personal con cuatro CPUs, por lo que en vez de poder correr 30 o más casos a la vez, estos han ido ejecutándose de cuatro en cuatro, lo que ha provocado que el factor tiempo sea un limitante en el momento de realizar este proyecto. Debido a ese factor tiempo, en lugar de ejecutar todos los casos posibles, se han terminado ejecutando 4000, por lo que se ha de comprobar que este número de casos es suficiente para que la red aprenda de forma correcta los patrones y características significativas. En el apartado 8 se mostrarán los resultados y se verá si con los casos ejecutados las predicciones que se realizarán se acercan a la realidad.

Tras la ejecución de los casos, se hace nuevamente uso de Python para desarrollar un script que obtenga la información de interés de los archivos de salida y la disponga de tal manera que facilite el tratamiento de datos posterior que se llevará a cabo cuando se diseñe la red neuronal.

### 7.4 Datos obtenidos

Para poder generar un número de inputs lo suficientemente grande, los datos se han obtenido de forma aleatoria salvo el quemado, ya que se han eliminado del set de datos todos aquellos casos en los que el quemado no se encontraba entre 38000 y 50000 MWd/tU. Al tratarse de datos aleatorios, si bien cada uno de ellos está dentro de un rango que podría darse en operación, su combinación de manera aleatoria puede llevar a resultados alejados de la realidad en la operación del combustible en el reactor. La limitación en los valores del quemado trata de minimizar el número de casos en los que se produce este efecto.

Como ejemplo de lo anterior, a pesar de que los valores de los periodos de irradiación y las potencias específicas son los más habituales, una mala combinación puede dar lugar a problemas. Esto es debido a que en la realidad hay que ir adaptando el combustible a los quemados, de tal forma que no puede ocurrir que se queme a baja potencia en los primeros ciclos y a una potencia muy alta en el último, el resultado puede llevar a problemas con la termohidráulica del combustible.

Anteriormente se ha indicado que algunas de las barras de los elementos combustibles incorporan gadolinio junto con  $\text{UO}_2$ . En la Tabla 3 se compara la evolución en la constante de multiplicación efectiva en los casos de elementos combustibles con y sin gadolinio para un mismo elemento, con los mismos ciclos de quemado. Se puede observar claramente el efecto que tiene este elemento en reducir la reactividad inicial del combustible fresco reduciendo el valor de la  $K_{\text{eff}}$  con respecto al otro caso. A medida que se va quemando el gadolinio, el efecto que tendrá será menor hasta el punto en el que la reactividad del combustible baje lo suficiente o el gadolinio se queme por completo.

*Tabla 3: Comparación de caso con y sin gadolinio*

Tiempo (días)	K-eff sin gadolinio	K-eff con gadolinio
0	1,37157	1,04693
180	1,34087	1,08579
540	1,27521	1,15106
930	1,23458	1,2217
1320	1,20901	1,20513
1680	1,17668	1,17399
2070	1,13989	1,13833
2430	1,09939	1,09913
68023	0,96879	0,9665

A pesar de que hay algunos casos que tratan situaciones alejadas de las de operación real que se dan en un reactor, no se puede analizar cada caso uno a uno puesto que la cantidad de casos simulados es muy elevada y el coste temporal que supondría dicho análisis es enorme. Además, el hecho de que algunos casos simulados se alejen un poco de la operación normal implica que la red neuronal tendrá un abanico de datos mucho más amplio de donde poder identificar patrones y características singulares.

## 7.5 Desarrollo de la red neuronal

Una vez generados los datos de entrada necesarios, se puede dar comienzo a la segunda fase.

La red neuronal que se va a desarrollar a lo largo de este trabajo se hará mediante el lenguaje de programación Python utilizando la librería Tensorflow, la cual es una librería de código libre para aprendizaje automático desarrollada por Google. Se trata de una herramienta que ayuda a gestionar los procesos de aprendizaje, permitiendo el entrenamiento e implementación de un modelo de forma sencilla.

A la hora de desarrollar una red neuronal, para hacerlo de forma correcta, se seguirá el ciclo de vida de las redes neuronales. El ciclo comienza con la definición del modelo, después, se deberá compilar para que posteriormente se pueda entrenar y evaluar. Una vez realizados los pasos anteriores, el modelo será capaz de realizar predicciones.

Para poder llevar a cabo el trabajo de forma correcta, es necesario tener en cuenta el tipo de problema que se quiere resolver con la red neuronal, debido a que no se diseña igual una red para un problema de regresión, donde el objetivo es calcular la salida numérica de un valor, o de clasificación, donde el objetivo es seleccionar una clase dentro de una lista de clases. En este trabajo, para el diseño de la red neuronal se han seguido los siguientes pasos.

### Preparar los datos

Probablemente lo más importante cuando se trata de diseñar una red neuronal son los datos, puesto que la red se nutre de ellos para entrenarse, y es esta parte de preprocesado de datos a la que hay que dedicar mayor cantidad de tiempo.

Para que la red aprenda una serie de patrones de manera que, en un futuro, pueda realizar predicciones adecuadas, se necesitan preparar los datos. En primer lugar, se deben limpiar los datos, eliminando cualquier dato que pueda estar duplicado y aquellos que tengan alguna de sus características en blanco.

Una vez se tienen los datos limpios, se debe dividir el set de datos en dos grupos, entrenamiento y prueba. En este caso el 90% de los datos están destinados al entrenamiento mientras que el 10% restante se utilizarán para comprobar la precisión del modelo.

Para que el modelo entrene de forma efectiva, hay que alimentar durante el entrenamiento tanto las características, o datos de entrada, como las salidas, de manera que pueda ir identificando patrones y características significativas. Sin embargo, cuando se pone a prueba el modelo, los datos de salida solamente los utiliza para comparar dicho valor con el obtenido al realizar la predicción, y, posteriormente, mostrar el error cometido.

A menudo, las unidades de los distintos datos de entrada que se introducen provocan que el modelo dé mayor importancia a aspectos que pueden ser menos influyentes en el resultado, mientras que otros aspectos con mayor influencia se pueden dejar un poco de lado. Por esta razón, para que las unidades no influyan en el modelo, es muy útil normalizar los datos de entrada. Para poder realizar este proceso, Tensorflow dispone una función que permite observar las propiedades estadísticas de los datos de entrada mostrando aspectos como la media, desviación estándar o máximos y mínimos entre otros. Con el objetivo de normalizar, a cada valor de los datos de entrada se le restará la media y se dividirá entre la desviación estándar.

### Definir el modelo

El problema en cuestión a resolver en este trabajo es un problema de regresión, teniendo como objetivo calcular la  $K_{eff}$  y las concentraciones de U-235, Pu-239 y Am-241. Como se trata de predecir un valor numérico y no asignar a cada valor una categoría, y que los datos van a estar en formato tabla, el tipo de red más adecuado va a ser un MLP. En este modelo, cada nodo de una capa está conectado a todos los nodos de la siguiente.

Como modelo de la red diseñada en este trabajo, se ha optado por uno secuencial, definiendo una capa tras otra. Si bien el modelo funcional es más avanzado, al tratarse de construir una red neuronal desde cero y al tratarse de un problema de regresión, el modelo secuencial ofrece un mayor entendimiento de la red, junto con la capacidad de investigar entre distintas configuraciones de capas y de neuronas de forma más sencilla.

Dentro de la arquitectura de la red, no hay un método exacto para saber cuántas capas o cuántas neuronas por capa se deben colocar, por ello se va a seguir un proceso de prueba-error buscando minimizar el error tanto en el entrenamiento como en las pruebas, de manera que las predicciones realizadas se acerquen lo máximo posible a una simulación con SCALE.

En cada capa de neuronas es necesario definir una función de activación de manera que se permita que haya relaciones no lineales. Teniendo en cuenta las condiciones a cumplir por una función para poder ser utilizada como función de activación en una red neuronal, las más adecuadas son las Rectified Linear Unit (ReLU) o alguna de sus variaciones. Estas funciones solamente se pueden emplear en las capas ocultas, por lo que en la capa de salida las relaciones entre entradas y salidas serán lineales al carecer de función de activación.

## Compilar el modelo

Para compilar el modelo, se debe seleccionar una función de coste y un algoritmo de optimización de manera que, posteriormente, durante el entrenamiento la red trate de disminuir dicha función y el proceso se realice de forma óptima.

Para la función de coste, se ha seleccionado el error absoluto medio, pero también se monitorizará el error mínimo cuadrático, ambos son las métricas más habituales en problemas de regresión. En el caso del algoritmo de optimización, los adecuados para este tipo de problemas son el RMSProp y el Adam, puesto que ambos dan buenos resultados en problemas de regresión. El optimizador se encarga de calcular el gradiente de la función de coste por cada peso y, para minimizar el error, modifica cada peso en la dirección negativa del gradiente.

## Ajustar el modelo

Tras compilar el modelo y seleccionar los parámetros necesarios, llega el momento del entrenamiento. Este se consigue mediante una sola línea de código en la que se concretan los parámetros necesarios.

En primer lugar, se seleccionan los datos de las características, correspondientes a los valores de entrada, y los valores objetivo, los cuales son el resultado obtenido de las simulaciones. Después, se especifica el número de épocas, es decir, el número de veces que la red va a ver cada ejemplo de entrenamiento. En un principio, puede parecer que cuanto mayor sea el número de épocas, mejor será el entrenamiento, pero no es del todo cierto. Con un mayor número de épocas, sí que el error de entrenamiento se verá disminuido cada vez más, pero, por otro lado, la red neuronal, en vez de aprender a reconocer patrones y características significativas, está memorizando los datos y a pesar de que el error de entrenamiento baje, el error en la prueba subirá. Esto es lo que se conoce como sobreajuste u overfitting. Es por ello que no se puede saber un número concreto de épocas a establecer a priori, sino que también se requiere realizar pruebas para determinarlo.

Otro parámetro, opcional pero recomendable, es el del tamaño de lote, que permite separar el set de datos en paquetes más pequeños para facilitar el entrenamiento de la red. Junto a este parámetro, se puede asignar también un porcentaje de validación de manera que se reserve una serie de datos para probar el modelo.

Por último, se puede añadir otro parámetro para que, mientras dure el entrenamiento, se vaya mostrando información. Este se puede ajustar para mostrar la información que se desee, o incluso no mostrar nada y que aparezca solo el error al final.

## Evaluar el modelo

Una vez realizado el entrenamiento de la red, llega el momento de evaluar el modelo. Esto se realiza con datos que la red no ha visto antes y se pueden comparar las predicciones que ha realizado el modelo con los resultados obtenidos a través de la simulación.

Con el objetivo de visualizar el entrenamiento y la evaluación de mejor forma, se ha incluido en la red neuronal la realización de unos gráficos con el objetivo de mostrar la evolución del error en función de las épocas y cómo de lejos están los valores predichos de los reales.



## Hacer predicciones

Tras completar todos los pasos anteriores, el modelo ya estaría listo para hacer predicciones, que es el objetivo final del diseño de la red neuronal. Para ello, se debe insertar un tensor con los datos de entrada referentes al caso que se quiera predecir y normalizarlos, de manera que sean del mismo tipo que los datos utilizados para el entrenamiento y sirvan los patrones aprendidos en este.

## 8. Resultados

En este apartado se presentarán los resultados obtenidos tras el entrenamiento de la red neuronal. El número final de casos ejecutados en el programa SCALE ha sido de 4000 por motivos de tiempo, y se ha reservado un 10% de los casos para evaluar el modelo, por lo que el set de datos disponible para realizar el entrenamiento ha sido de dicho tamaño. Con esta cantidad de datos se tratará de ver si el error cometido en las predicciones es razonablemente pequeño, además, se comparará entre las predicciones realizadas con 3000 casos ejecutados con las realizadas tras la ejecución de 4000 casos para observar si un aumento del set de datos provoca una disminución significativa en el error. Para ver cuál será la red adecuada, debido a que no existe un método perfecto, se ha llevado a cabo un proceso de prueba error.

A la hora de diseñar la red se han realizado pruebas con 2 y 3 capas, dentro de ellas, se han realizado combinaciones con 64, 128 y 256 neuronas por capa. Se han seleccionado estas cantidades puesto que un número de capas ocultas mayor que 3 es más propenso al problema del sobreajuste y ralentiza en exceso el proceso de entrenamiento. Por otro lado, el componer una capa con una gran cantidad de neuronas no es adecuado debido a que el tamaño del set de datos de entrenamiento no es muy grande, por lo que aumentar este número puede llevar a la red a reducir su capacidad de distinguir patrones concretos y hacer predicciones precisas. También se ha experimentado con la función de activación de las neuronas, haciendo pruebas con la función ReLU y su variante PReLU. Es necesario aclarar que todas las pruebas se han realizado con el set de datos reducido, con 3000 casos ejecutados, es posteriormente, cuando ya se ha elegido la red adecuada, el momento en el que se ha vuelto a entrenar con el set de datos ampliado.

Durante la realización de las pruebas se ha monitorizado el error absoluto medio y el error cuadrado medio de manera que para la selección del modelo adecuado se tenga en cuenta el mejor en estos aspectos, teniendo en cuenta que en cualquier prueba realizada el tiempo de entrenamiento no se ha considerado excesivo.

Se han realizado dos pruebas distintas para cada arquitectura y función de activación distinta, se hace de esta manera porque al comienzo, los parámetros se inicializan de forma aleatoria y hay una pequeña probabilidad de que tras el entrenamiento se pueda obtener un valor del error incoherente. Se realiza la segunda prueba para asegurar que el valor del error obtenido es bueno. Los resultados de la red global se muestran en la tabla 4, donde en los valores de los errores se encuentra una media de las dos pruebas realizadas en cada caso.

Tabla 4: Resultados de la red global

ARQUITECTURA	RED GLOBAL			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 64	28,00	3950,00	36,70	9990,00
128 - 64	25,00	3200,00	24,50	5430,00
64 - 128	47,50	8800,00	18,20	838,00
128 - 128	24,50	2797,00	17,00	712,00
256 - 128	21,00	1768,50	16,00	810,00
128 - 256	23,00	1783,00	16,00	745,00
256 - 256	21,80	1849,50	13,40	425,00
64 - 64 - 64	25,00	3116,50	18,85	903,50
128 - 64 - 64	20,65	1476,50	16,85	687,00

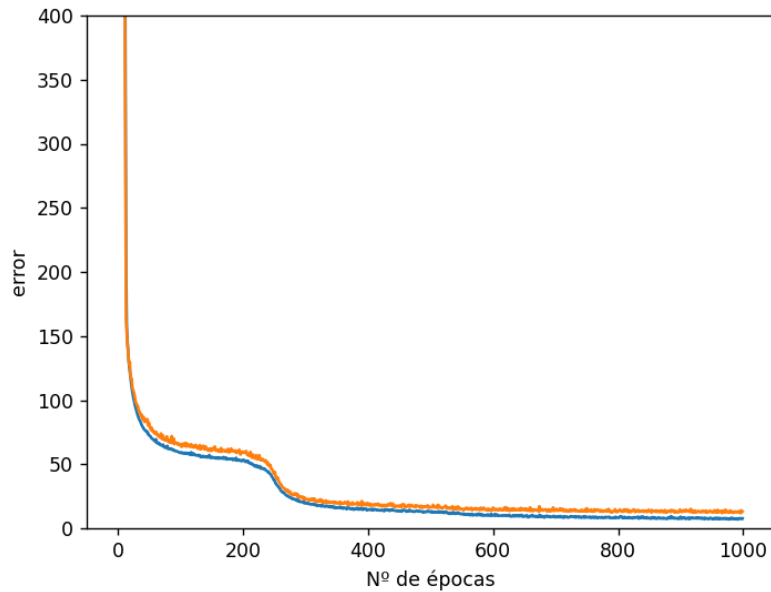
ARQUITECTURA	RED GLOBAL			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 128 - 64	38,75	5508,00	17,65	761,50
64 - 64 - 128	49,50	7588,50	19,95	962,50
128 - 128 - 64	23,35	2124,00	16,60	967,50
128 - 64 - 128	25,10	2103,00	17,20	718,00
64 - 128 - 128	21,84	1362,50	16,95	809,50
128 - 128 - 128	20,85	1245,50	16,75	654,50
256 - 128 - 128	19,05	1025,50	14,95	536,50
128 - 256 - 128	20,20	1115,00	16,55	669,00
128 - 128 - 256	17,60	830,50	15,60	574,50
256 - 256 - 128	19,70	910,00	14,10	510,50
256 - 128 - 256	17,60	768,50	18,00	752,50
128 - 256 - 256	18,95	857,50	15,55	624,00
256 - 256 - 256	20,75	1047,50	16,60	697,50

Tras el análisis de los resultados expuestos en la tabla, se puede sacar información muy interesante. En primer lugar, para este caso, salvo en algún caso en concreto, la función de activación PReLU funciona mucho mejor que la ReLU. También hay que destacar los casos en los que el número de capas y neuronas es el mismo, donde los errores difieren, esto hace entender la importancia de la arquitectura y que la forma en la que están dispuestas las capas puede tener una importancia significativa para identificar patrones a pesar de tener un número similar de neuronas.

Algo más que se puede decir de estos resultados es que, por lo que se ve en el caso de las 3 capas, el error no sigue necesariamente una tendencia parabólica, donde va descendiendo hasta llegar a un mínimo y a partir de ahí irá creciendo debido al sobreajuste.

Además de esta información, lo importante de estos resultados es hallar el modelo que realizará las mejores predicciones, en este caso se trata de la arquitectura de dos capas con 256 neuronas por capa. El error absoluto medio en esta es de 13, este carece de unidades debido a que evalúa el error correspondiente a las 4 predicciones y los une. Para obtener ese valor, la red obtiene las salidas de los 4 vectores, a cada una de ellas se le resta el valor real y, por último se realiza la suma de todos los valores y se divide entre el número total de componentes. Al mezclarse en este caso predicciones correspondientes a concentraciones de isótopos y la  $K_{eff}$ , el error no tendrá unidades.

En la ilustración 26 se muestra la evolución de error a lo largo de las distintas épocas. Ahí se aprecia que el error sufre una brutal caída en las primeras épocas debido a la inicialización aleatoria de los pesos y sesgos de la red unido al proceso de entrenamiento. En este caso el error en el entrenamiento (línea azul) y el error en la prueba (línea naranja) son muy parecidos, y al no crecer este último, se puede decir que no se produce sobreajuste. A partir de las 600 épocas los errores son prácticamente horizontales por lo que el entrenamiento se podría haber terminado antes. De haberlo hecho se hubiese compensado un ligero aumento en el error con una reducción en el tiempo de entrenamiento.



*Ilustración 26: Error en función del número de épocas en la red global*

Tras el análisis de la evolución del error global, se analiza el rendimiento de la red a la hora de predecir cada uno de los parámetros comparando el valor de la predicción con el valor real de salida. en las ilustraciones que se presentan a continuación se representan con puntos azules los valores de las predicciones (eje de ordenadas izquierdo) frente a los valores reales de salida (eje de abscisas), mientras que se dibuja también una línea naranja la cual es una recta de ecuación  $y=x$  que representa la predicción ideal, cuanto más se acerquen los puntos a la recta, mejor será la predicción y menor el error. Además, se muestra con los puntos verdes el error absoluto (eje de ordenadas derecho).

En primer lugar, en la ilustración 27 se muestran las predicciones de uranio. Por lo que se muestra en la gráfica, la red es perfectamente capaz de realizar unas predicciones más que aceptables, aunque también es cierto que, al ser la escala de valores tan grande, un error de 100 unidades de diferencia entre predicción y valor real sería prácticamente inapreciable en la imagen. Sin embargo, también es cierto que, al estar la media en unos 6500 gramos, ese error hipotético de 100 unidades supondría un error menor del 5%.

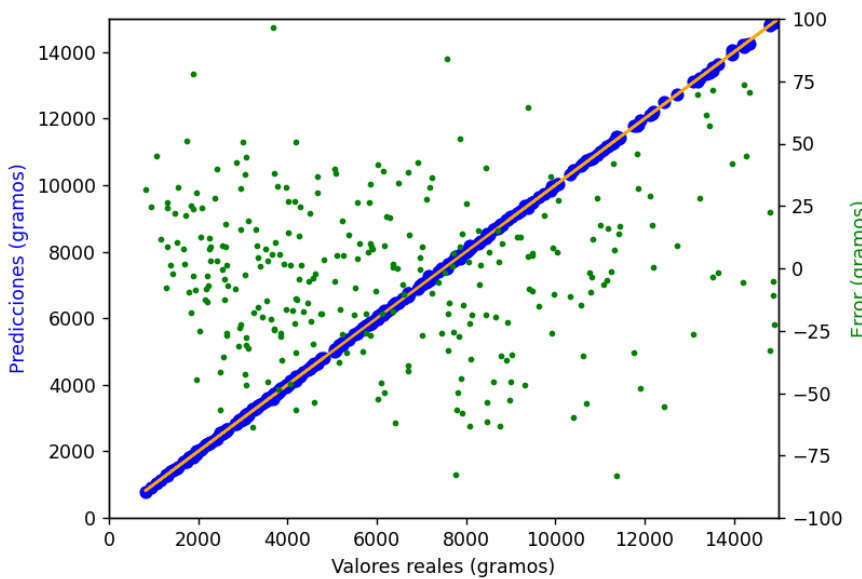


Ilustración 27: Predicciones frente a valores reales del U-235 en la red global

Siguiendo con los valores del plutonio, se observa en la ilustración 28 que las predicciones también son bastante buenas, en un principio puede parecer que el error está menos disperso, pudiendo deberse a que los valores son menores. Por otro lado, comparando con el caso del uranio, puede parecer que las predicciones se alejan un poco más del valor real de salida obtenido mediante simulación. En este parámetro, el valor medio está en torno a los 5400 gramos, solo que el rango de valores obtenido es menor que en el caso del uranio.

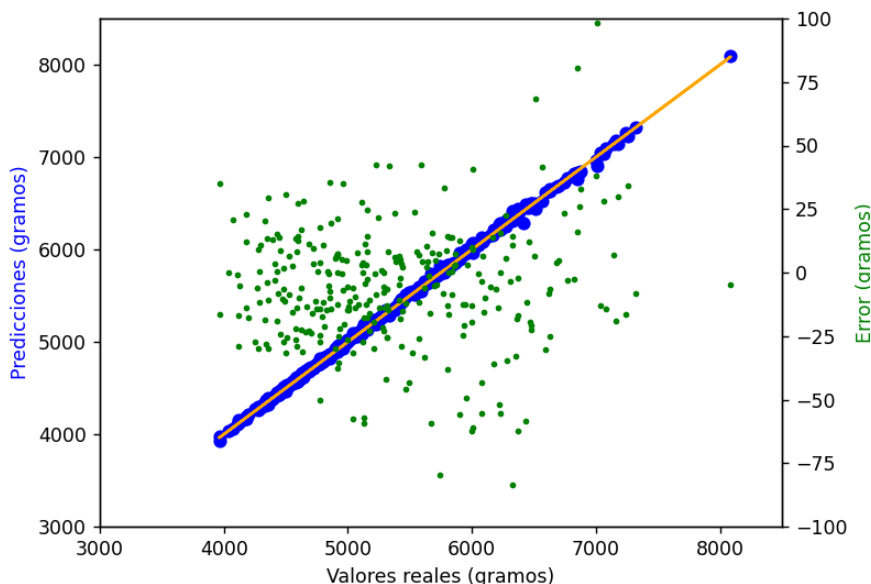


Ilustración 28: Predicciones frente a valores reales del Pu-239 en la red global

En el caso del americio, se observa que hay ciertos valores de las predicciones que sí se alejan bastante de los valores de salida, sobre todo en los casos en los que se obtienen cantidades pequeñas de este isótopo. Se ve que hay errores de incluso unos 200 gramos, oscilando los valores de americio entre 150 y 1600 gramos, con una media de 950.

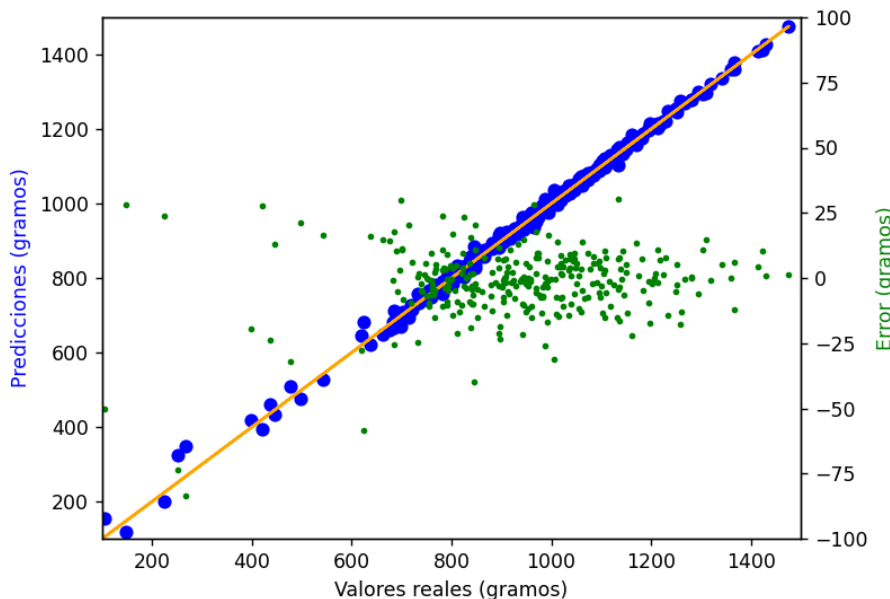


Ilustración 29: Predicciones frente a valores reales del Am-241 en la red global

Por último, se muestra en la ilustración 30 las predicciones referentes a la constante de multiplicación efectiva, y se puede apreciar perfectamente que dichas predicciones para este parámetro son nefastas, no consiguiendo predecir ningún valor de forma correcta. En este caso, un valor de la  $K_{eff}$  mayor que 1 significa criticidad, por lo que se pueden dar fisiones, cosa que hay que evitar a toda costa. Teniendo esto en cuenta, las predicciones realizadas indican que todos los valores son muy superiores a 1, valores que no se pueden obtener en operación y mucho menos durante el almacenamiento de combustible gastado. Además, solamente se puede ver un punto verde, por lo que el error está muy fuera de la escala.

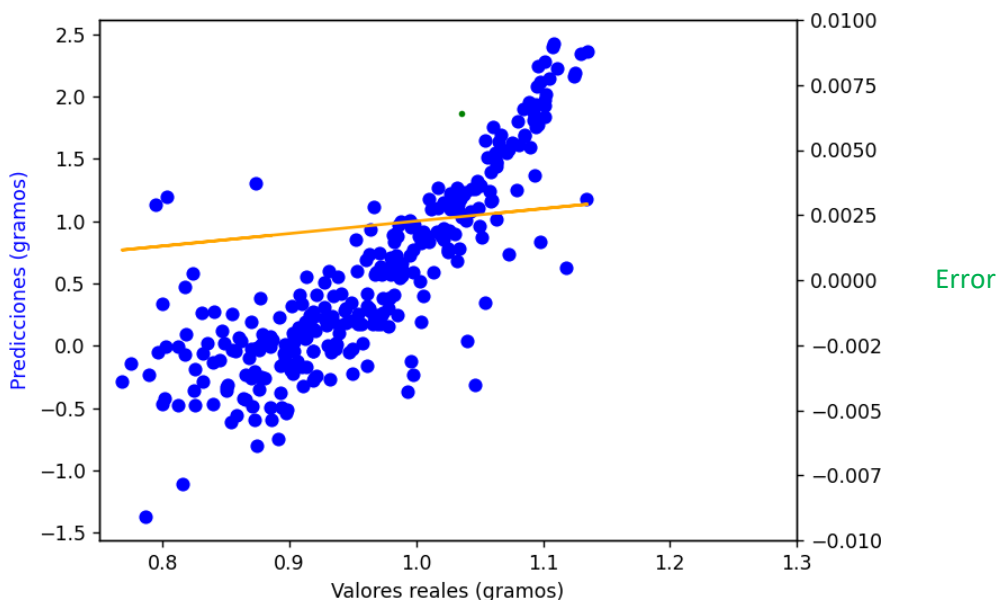


Ilustración 30: Predicciones frente a valores reales de la  $K_{eff}$  en la red global

Vistos los resultados, se puede decir que las predicciones son bastante mejorables, una de las razones que existen para explicar los errores que se dan es que para cada parámetro hay que tener en cuenta diferentes características y, al estar prediciendo todo en una misma red, no es posible modificar los pesos y sesgos en función del parámetro a predecir, alcanzando un valor final de ellos que se acerca más a la identificación de

patrones y características significativas relativas al uranio o plutonio, ya que son los que mejor resultados han obtenido.

Viendo los resultados obtenidos tras la ejecución de 3000 casos, con la misma arquitectura escogida se ha vuelto a entrenar la red, pero aumentando el set de datos. La siguiente tabla compara los resultados obtenidos con el set de datos anterior y el ampliado.

Tabla 5: Comparación de los errores en los set de datos reducido y ampliado

SET DE DATOS	EMA	EMC
3000 casos	13,4	425
4000 casos	11,5	300

Se ve perfectamente que el error es menor, sin embargo, hay que analizar también cada caso particular para ver si existen mejoras en las predicciones realizadas.

La ilustración 31 muestra el caso del uranio, donde, comparando con las predicciones realizadas con un menor número de casos ejecutados no se ve mucha diferencia. Esto puede ser porque, como se ha dicho previamente, la escala es muy grande y no se aprecia muy bien el tamaño de los errores.

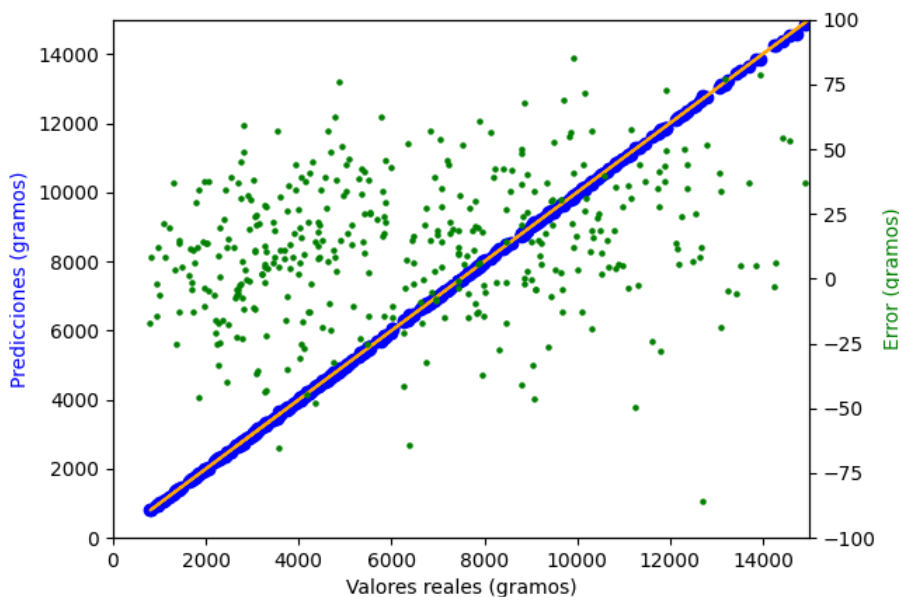


Ilustración 31: Predicciones frente a valores reales del U-235 en la red global con 4000 casos

En el caso del plutonio ocurre lo mismo, las predicciones también parecen bastante buenas, pero no se puede comparar debido a que no se aprecia muy bien el tamaño real del error.

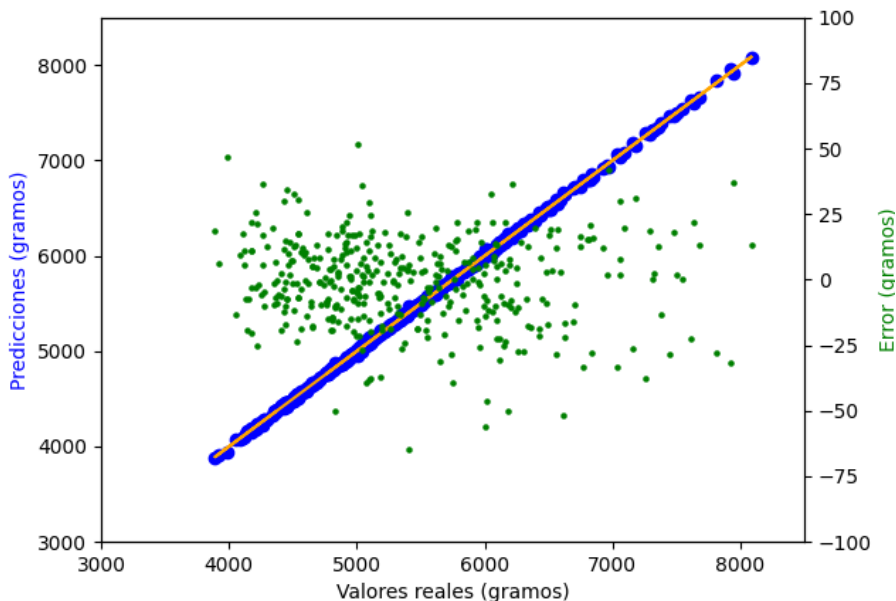


Ilustración 32: Predicciones frente a valores reales del Pu-239 en la red global con 4000 casos

Tras realizar las predicciones en el americio, en la ilustración 33 sí que se puede observar cierta mejora en los resultados de las predicciones, sobre todo a valores bajos de concentración del isótopo. Es en esta categoría donde, por el momento, sí que se puede confirmar que aumentar el tamaño del set de datos mejora los resultados.

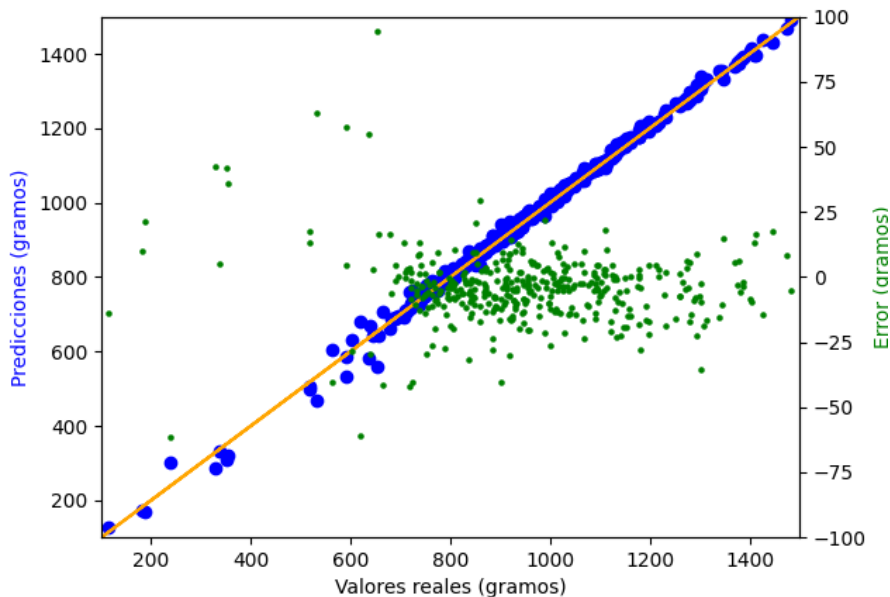


Ilustración 33: Predicciones frente a valores reales del Am-241 en la red global con 4000 casos

Por último, en el caso de la constante de multiplicación efectiva, las predicciones son igual de inservibles que en el caso con el set de datos reducido, por lo que tampoco se puede apreciar el efecto de aumentar el set de datos.



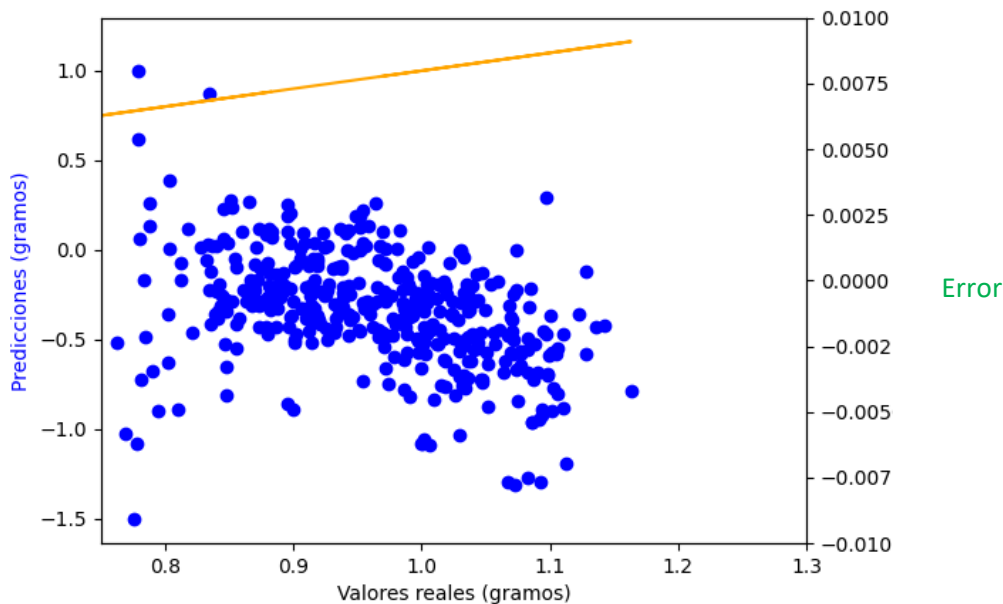


Ilustración 34: Predicciones frente a valores reales de la  $K_{eff}$  en la red global con 4000 casos

Vistos los resultados obtenidos en la red global al tratar de predecir los 4 parámetros al mismo tiempo, se ha decidido establecer una red neuronal para cada parámetro a predecir.

### 8.1 U-235

En primer lugar, se diseñará la red para la predicción de la concentración de U-235. El proceso que se ha llevado a cabo para establecer la red es el mismo que se ha seguido para la red global. La diferencia radica en que, en vez de cuatro salidas, solamente se guarda como salida el parámetro de interés de la red, en este caso las concentraciones de uranio.

Los resultados de las pruebas realizadas se recogen en la Tabla 6, donde esta vez el error absoluto medio sí que tiene unidades, los gramos.

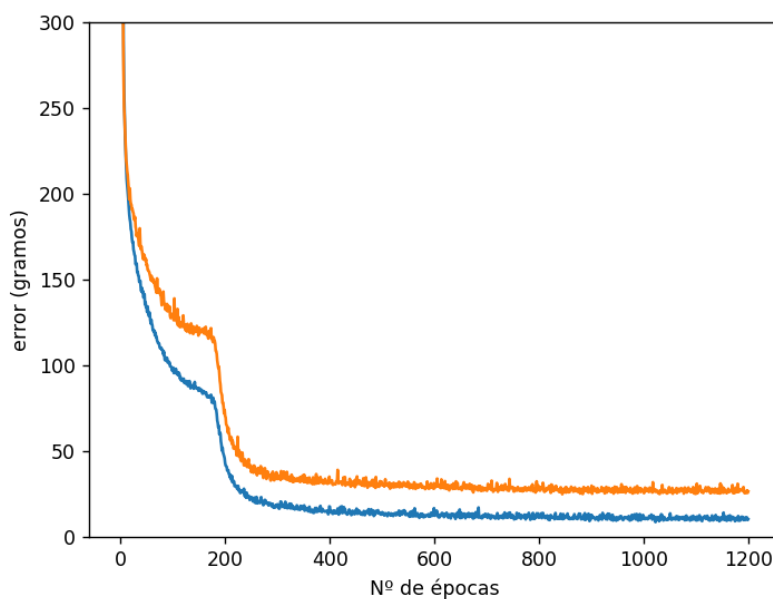
Tabla 6: Resultados de la red de U-235

ARQUITECTURA	CONCENTRACIÓN U-235			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 64	48,00	3731,50	31,65	1682,50
128 - 64	38,50	2440,50	26,70	1163,00
64 - 128	42,00	2904,50	32,60	1776,00
128 - 128	46,00	3536,50	31,75	1543,50
256 - 128	43,00	3107,50	28,65	1349,50
128 - 256	47,50	3835,50	25,65	1081,50
256 - 256	41,25	2968,50	25,50	1068,00
64 - 64 - 64	37,75	2422,00	30,40	1496,00
128 - 64 - 64	42,15	2933,50	33,50	1825,50
64 - 128 - 64	40,75	2577,00	31,60	1601,50
64 - 64 - 128	37,95	2338,00	33,05	1773,00
128 - 128 - 64	53,00	4579,00	32,40	1808,00
128 - 64 - 128	45,85	3331,50	31,95	1666,00

ARQUITECTURA	CONCENTRACIÓN U-235			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 128 - 128	44,15	3089,50	35,30	1933,50
128 - 128 - 128	48,20	3718,50	42,80	2970,00
256 - 128 - 128	37,60	2558,00	35,10	2074,50
128 - 256 - 128	54,00	4739,50	41,90	2770,00
128 - 128 - 256	49,70	4304,50	35,55	2045,50
256 - 256 - 128	40,55	2638,50	33,20	1853,00
256 - 128 - 256	40,20	2589,50	39,65	2562,00
128 - 256 - 256	53,65	4903,00	39,35	2562,00
256 - 256 - 256	50,40	4200,50	37,20	2195,00

En este caso, se vuelve a dar la situación donde la red con la función de activación PReLU trabaja mejor que con la función ReLU. Siguiendo también en la misma línea que la red global, el modelo de dos capas ofrece mejores resultados que el de tres, siendo también la arquitectura de dos capas de 256 la mejor para el diseño de la red. Este resultado casa con lo anterior debido a que en la red global el uranio ofrece muy buenos resultados, por lo que la estructura de la red debería ser parecida.

Echando un vistazo a la ilustración 35, donde se muestra la evolución tanto del error de entrenamiento (línea azul) como del error en la prueba (línea naranja), se ve que ahora estos no están tan unidos como antes. Por otro lado, a partir de las 600 épocas, los errores pasan a ser prácticamente horizontales, por lo que se podría haber detenido el entrenamiento antes. Sin embargo, en las 200 épocas el error parece que se va a volver horizontal, pero hubiera sido un error detener el entrenamiento tan pronto debido a que posteriormente sufre una caída brutal. Además, al no tener un crecimiento del error en la prueba, se puede concluir también que no se da sobreajuste.



*Ilustración 35: Gráfica del error frente al número de épocas*

Observando ahora los valores de las predicciones frente a los valores reales de salida de la ilustración 36, se ve que las predicciones son bastante buenas, igual que en la red global. Ahora, se sabe que el error

absoluto medio es de 25,5 gramos, algo que no se podía saber mirando la red global ya que mostraba el error unificado en los cuatro parámetros y no se podía observar uno en concreto. Teniendo en cuenta dicho error medio, y que la media de valores es de 6500 gramos, se puede afirmar que los resultados son óptimos.

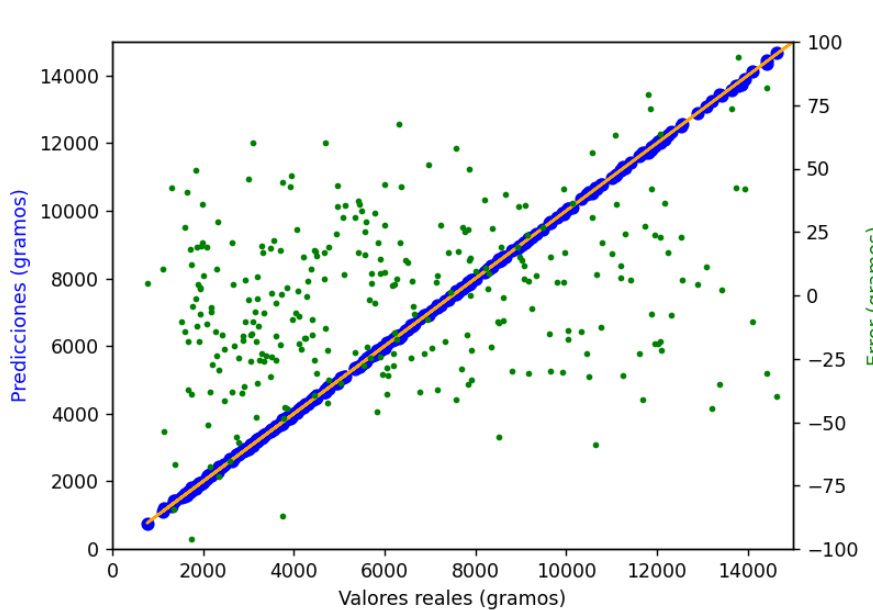


Ilustración 36: Predicciones frente a valores reales del U-235

Tras ampliar el set de datos se ha vuelto a entrenar la red neuronal, obteniendo una reducción importante en el error como se puede apreciar en la Tabla 7. Al tener el error individualizado para el caso concreto del U-235 sí que se puede afirmar que existe mejora, en oposición al caso de la red global, donde a pesar de que el error sí que es menor, no se puede dividir en las cuatro categorías.

Tabla 7: Comparación de los errores en los set de datos reducido y ampliado

SET DE DATOS	EMA	EMC
3000 casos	25,5	1068
4000 casos	19,9	630

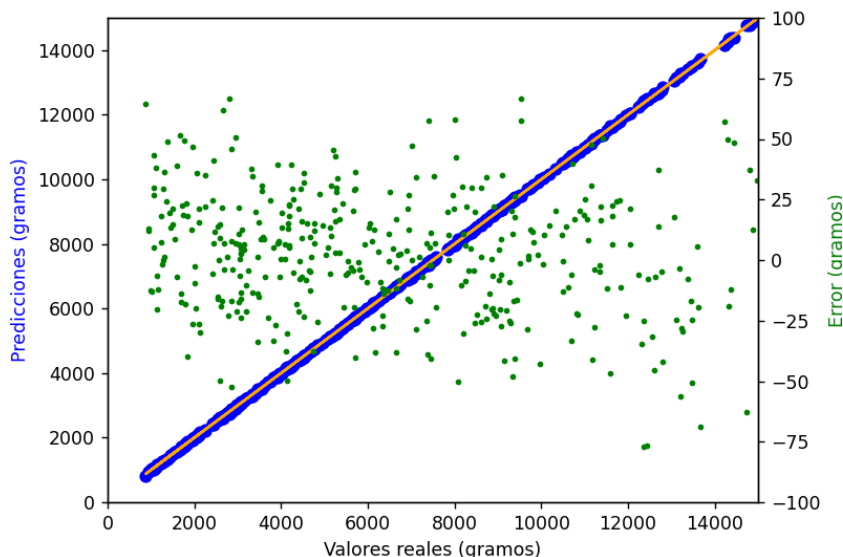


Ilustración 37: Predicciones frente a valores reales del U-235 con el set de datos ampliado

Atendiendo a la ilustración 37 y comparándola con la ilustración 36, se concluye que el hecho de ampliar el set de datos hace que el error se acote. En este caso no hay ningún error que se acerque a los 100 gramos por lo que, sobre todo a valores bajos de U-235 el porcentaje de error será menor, y las predicciones estarán más ajustadas.

## 8.2 Pu-239

Tras la red de U-235, se diseña la red de Pu-239 y, de nuevo, se sigue el mismo proceso tanto para el diseño como para la realización de las pruebas que indicarán la mejor arquitectura posible para minimizar el error. Los resultados de dichas pruebas se muestran en la Tabla 8 y a partir de ahí se decidirá la adecuada.

Tabla 8: Resultados de la red de Pu-239

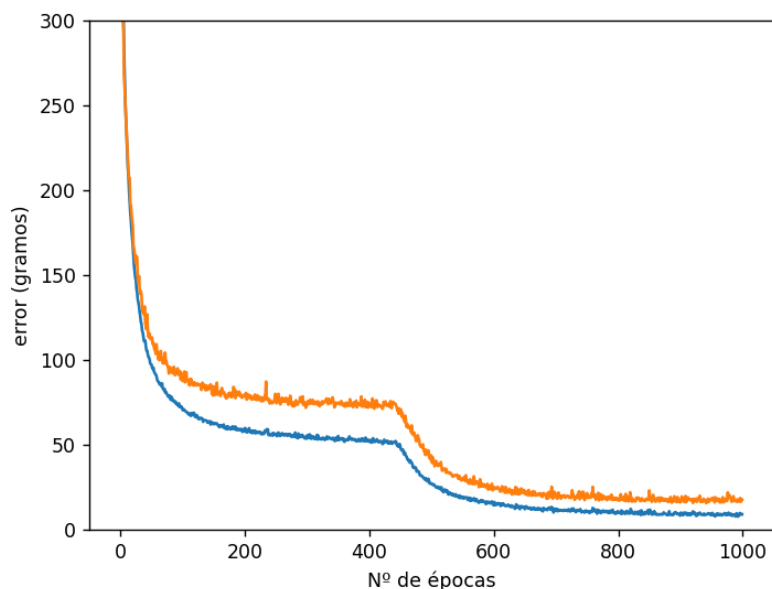
ARQUITECTURA	CONCENTRACIÓN Pu-239			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 64	22,00	803,00	21,25	784,50
128 - 64	19,95	701,00	18,05	574,00
64 - 128	24,15	980,50	20,95	811,00
128 - 128	17,20	475,00	17,35	519,50
256 - 128	16,85	518,50	16,30	446,50
128 - 256	15,60	419,00	16,05	450,50
256 - 256	15,10	370,00	16,25	468,50
64 - 64 - 64	21,25	737,50	21,05	767,50
128 - 64 - 64	19,10	603,50	17,10	521,00
64 - 128 - 64	29,45	2358,00	20,35	714,00
64 - 64 - 128	21,10	710,00	21,25	768,00
128 - 128 - 64	17,15	512,00	16,25	435,50
128 - 64 - 128	15,90	419,00	17,15	477,50
64 - 128 - 128	20,20	684,50	24,85	973,50
128 - 128 - 128	18,75	602,00	20,50	690,50
256 - 128 - 128	20,00	668,00	16,25	434,50

ARQUETECTURA	CONCENTRACIÓN Pu-239			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
128 - 256 - 128	17,95	516,00	17,30	513,00
128 - 128 - 256	26,50	1099,00	19,50	605,00
256 - 256 - 128	17,45	525,50	17,05	488,50
256 - 128 - 256	20,60	689,50	20,50	654,50
128 - 256 - 256	18,05	514,50	20,10	635,50
256 - 256 - 256	15,70	401,00	17,75	534,00

Viendo la ilustración 28 que muestra los valores predichos frente a los de salida de la red global, se ve que poco a poco los puntos se comienzan a separar de la línea de referencia, por lo que, a pesar de que la arquitectura de la propia red global no es del todo mala, puede que haya alguna que mejore los resultados. Se puede ver que en este caso tanto en el modelo de dos capas como en el de tres, no queda del todo claro que la función de activación PReLU sea mejor que la función ReLU, por lo que ya se empiezan a notar diferencias respecto a los dos casos anteriores. En este caso, y por tercera vez, la arquitectura que ofrece un menor error es la de dos capas con 256 neuronas por capa, sin embargo, ahora es la función ReLU la que ofrece un menor error.

Esto anterior explica que, aunque la misma arquitectura de la red global ofrece un error relativamente bajo, existe un diseño de red con mejores resultados.

Atendiendo a la ilustración 38, donde se muestra la evolución del error en función de las épocas, el error en la prueba vuelve a estar más cerca del error en el entrenamiento, alcanzando un valor prácticamente horizontal a partir de las 500 épocas. Se aprecia también un patrón común con los dos casos anteriores donde, alrededor de las 200 épocas se produce un amago de estabilización seguido de una bajada del error que no es tan pronunciada en este caso, pero sí se hace notar.



*Ilustración 38: Gráfica del error frente al número de épocas*

Si se observa la ilustración 39, se puede ver que se parece bastante a la ilustración 28 correspondiente a las predicciones de este isótopo de plutonio en la red global, debido a que esta ya realiza predicciones aceptables y su arquitectura es la misma. El único cambio de esta red respecto a la global es la función de activación y, la ilustración mostrada a continuación permite ver que, en este caso, utilizar la función ReLU resulta ligeramente mejor.

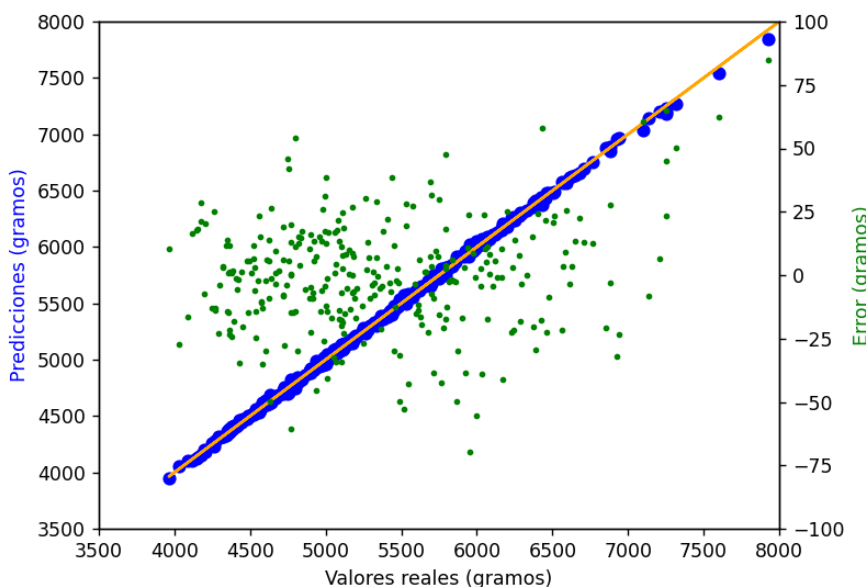


Ilustración 39: Predicciones frente a valores reales del Pu-239

Una vez ampliado el set de datos y vuelto a entrenar la red, el error vuelve a verse reducido, por lo que, si ya en el caso con el set de datos reducido las predicciones eran bastante buenas, ahora serán mejores. En la Tabla 9 se puede ver el error para cada uno de los casos.

Tabla 9: Comparación de los errores en los set de datos reducido y ampliado

SET DE DATOS	EMA	EMC
3000 casos	15,5	370
4000 casos	12,98	306

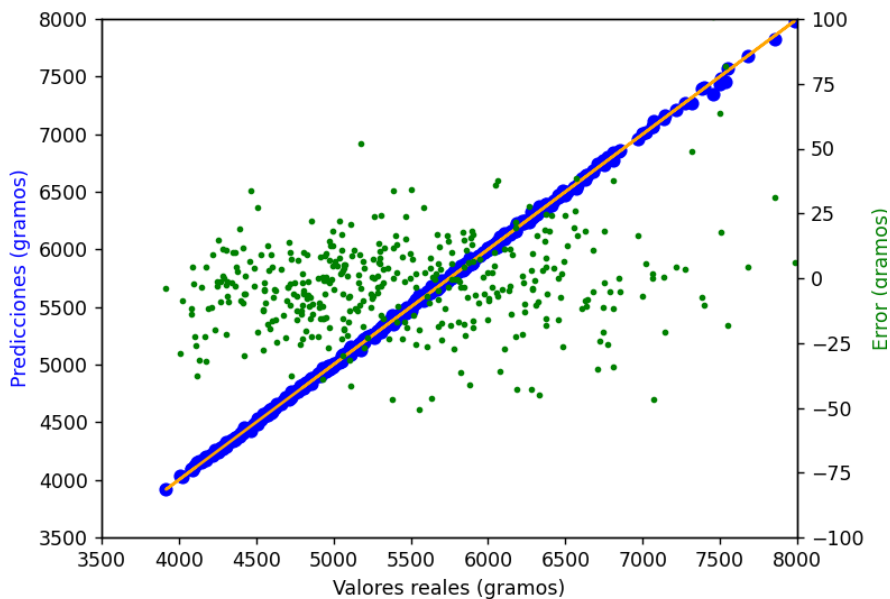


Ilustración 40: Predicciones frente a valores reales y error del Pu-239 con 4000 casos

Con el set de datos ampliado, atendiendo a la ilustración 40 se puede apreciar que los puntos verdes correspondientes a los errores están mucho más compactos en que con el set reducido, lo que significa una mejor predicción. Sin embargo, siguen existiendo valores relativamente altos, aunque no tanto como los que se ven en la ilustración 39.

### 8.3 Am-241

La siguiente red que diseñar es la del americio 241, siguiendo el mismo proceso y realizando las pruebas correspondientes se han obtenido una serie de resultados que se mostrarán en la Tabla 10 con el fin de escoger la mejor red posible.

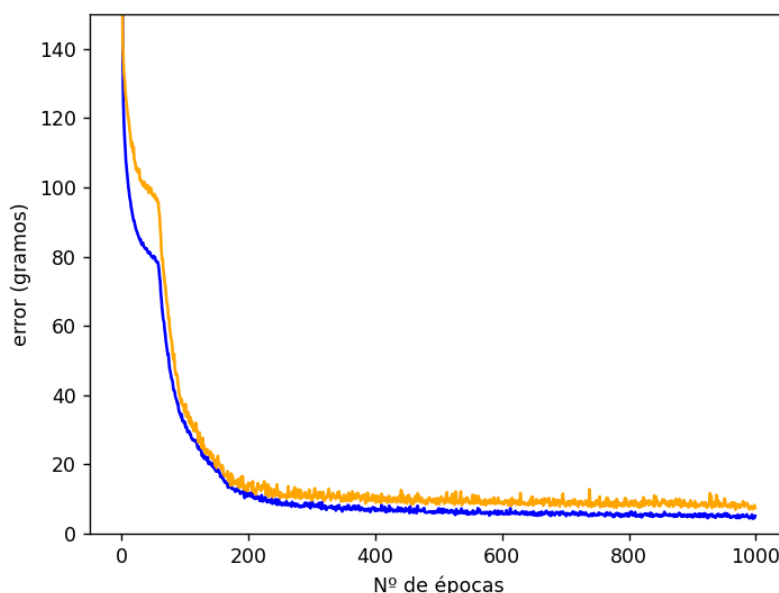
Tabla 10: Resultados de la red de Am-241

ARQUITECTURA	CONCENTRACIÓN Am-241			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 64	13,60	649,00	14,60	405,50
128 - 64	18,95	948,50	12,50	365,00
64 - 128	13,75	824,00	14,95	540,50
128 - 128	13,80	910,50	11,85	306,00
256 - 128	13,20	513,00	13,20	412,00
128 - 256	13,00	551,00	11,30	347,50
256 - 256	11,10	391,50	12,60	588,00
64 - 64 - 64	10,85	399,00	13,64	744,00
128 - 64 - 64	17,15	1121,00	9,70	301,00
64 - 128 - 64	9,35	200,85	7,80	286,00
64 - 64 - 128	9,80	270,50	9,45	268,00
128 - 128 - 64	15,05	720,50	12,10	658,50
128 - 64 - 128	13,10	663,50	12,15	562,50
64 - 128 - 128	12,40	559,50	9,55	294,50

ARQUITECTURA	CONCENTRACIÓN Am-241			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
128 - 128 - 128	13,50	539,00	9,35	261,00
256 - 128 - 128	14,60	764,50	11,40	555,00
128 - 256 - 128	20,60	1925,00	9,85	298,00
128 - 128 - 256	14,55	883,50	12,25	663,00
256 - 256 - 128	11,80	771,00	17,50	1994,50
256 - 128 - 256	9,90	524,50	12,35	519,00
128 - 256 - 256	14,15	658,00	12,75	725,50
256 - 256 - 256	9,60	352,50	13,90	1163,00

En este caso se puede ver la razón por la que en la red global las predicciones no eran tan buenas como para los otros parámetros, esta red ofrece mejores resultados con arquitecturas de 3 capas que con las de 2. Por otro lado, no queda del todo claro qué función de activación es mejor puesto que en ocasiones, la más adecuada es la ReLU, mientras que en otras es la PReLU. Con todo esto, se ha decidido escoger como arquitectura para esta red el modelo de tres capas con 64 neuronas en la primera capa, 128 en la segunda y 64 en la tercera.

En la ilustración 41 se vuelve a analizar la evolución tanto del error de entrenamiento como del error de prueba en función del número de épocas. Se puede apreciar que, en torno a las 150 épocas, del mismo modo que en casos anteriores, el error parece que comienza a seguir una tendencia horizontal, pero, poco más tarde sufre una gran caída para, ahora sí, estabilizarse sobre las 500 épocas.



*Ilustración 41: Gráfica del error frente al número de épocas*

Centrando las miradas en la ilustración 42, donde se representan las predicciones frente a los valores de salida reales, y comparándolos con la ilustración 29 correspondiente a este isótopo tratado en el caso global, se puede apreciar que, en la zona cercana a la media de valores, las predicciones son óptimas en



este caso. Cuando los valores de este isótopo de americio son bajos, el modelo parece que funciona peor y las predicciones se alejan un poco de lo que deberían ser.

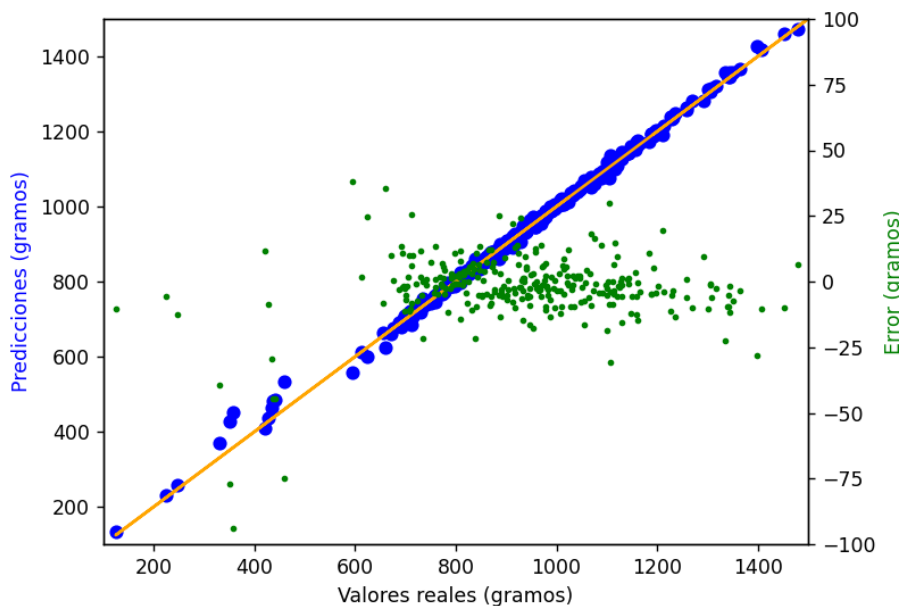


Ilustración 42: Predicciones frente a valores reales del Am-241

El aumento del set de datos vuelve a traer consigo una reducción del error, como muestra la Tabla 11, y esta vez, al contrario que en los casos anteriores, sí que se hace evidente por lo que muestra la ilustración 43 comparándola con la ilustración 42, sobre todo a valores bajos de la concentración de Am-241. Se ve que en general, los valores son mejores, a pesar de que haya algún caso aislado en el que el error es relativamente alto.

Tabla 11: Comparación de los errores en los set de datos reducido y ampliado

SET DE DATOS	EMA	EMC
3000 casos	7,8	286
4000 casos	6,8	120

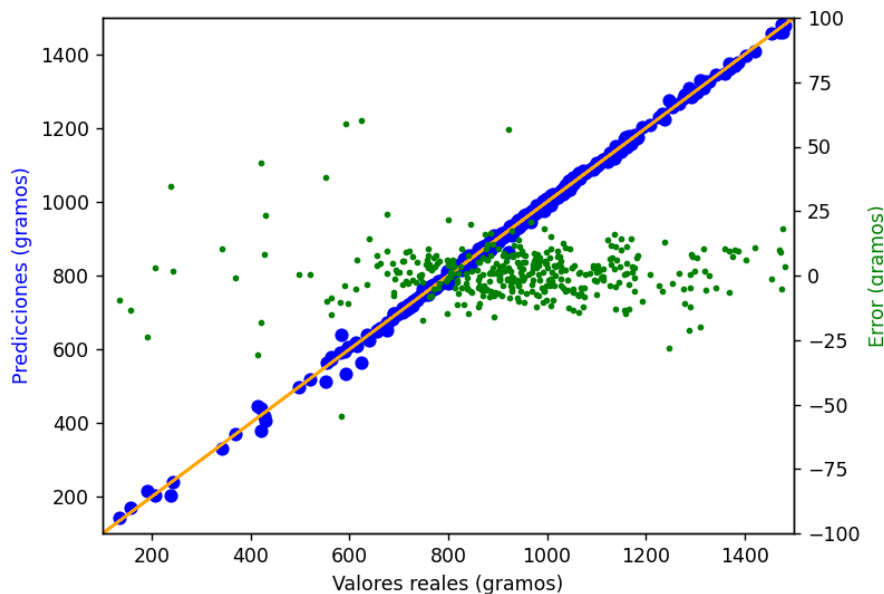


Ilustración 43: Predicciones frente a valores reales del Am-241 con 4000 casos

#### 8.4 Constante de multiplicación efectiva

La última red que queda por definir es la de la constante de multiplicación efectiva, la cual, en el caso de la red general era la que peores resultados ofrecía. Tras el diseño de la red y la realización de las pruebas pertinentes, los resultados obtenidos han sido los mostrados en la Tabla 12.

Tabla 12: Resultados de la red de la  $K_{eff}$

ARQUITECTURA	K-EFF			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
64 - 64	0,00200	0,85	0,00155	0,45
128 - 64	0,00195	0,85	0,00169	0,635
64 - 128	0,00180	0,67	0,00180	0,58
128 - 128	0,00200	0,9885	0,00200	0,725
256 - 128	0,00250	1,22	0,00163	0,45
128 - 256	0,00215	0,765	0,00185	0,53
256 - 256	0,00235	1,005	0,00200	0,925
64 - 64 - 64	0,00300	1,55	0,00210	1,075
128 - 64 - 64	0,00250	1,55	0,00150	0,535
64 - 128 - 64	0,00340	2,3	0,00140	3,6 E-05
64 - 64 - 128	0,00345	2,55	0,00170	0,515
128 - 128 - 64	0,00290	1,9	0,00242	1,295
128 - 64 - 128	0,00210	1,25	0,00165	0,57
64 - 128 - 128	0,00320	2,1	0,00160	0,625
128 - 128 - 128	0,00215	1,07	0,00200	0,88
256 - 128 - 128	0,00230	1,22	0,00159	0,485
128 - 256 - 128	0,00370	2,735	0,00165	0,63
128 - 128 - 256	0,00240	1,135	0,00200	0,86
256 - 256 - 128	0,00290	1,58	0,00185	0,72

ARQUITECTURA	K-EFF			
	Func. act. ReLU		Func. act. PReLU	
	EMA	EMC	EMA	EMC
256 - 128 - 256	0,00260	1,4	0,00220	1,125
128 - 256 - 256	0,00235	1,3	0,00235	1,12
256 - 256 - 256	0,00245	1,16	0,00185	0,98

Para este caso sí se puede decir que, para la mayoría de las arquitecturas probadas, la función de activación PReLU genera mejores resultados que la función ReLU. Observando la tabla superior, se concluye que, para este parámetro, son los modelos de 3 capas con menos neuronas por capa los que mejor funcionan, en contraste con el modelo de dos capas con un mayor número de neuronas. Esta razón puede que sea uno de los motivos principales por los que la red global ofrece resultados tan pobres a la hora de predecir este parámetro.

Tras el análisis de los resultados, se ha escogido como arquitectura para la red la de tres capas con 64 neuronas en la primera capa, 128 en la segunda y 64 en la tercera.

Atendiendo a la ilustración 44, se puede observar una tendencia distinta al resto de los casos. Para la predicción de la  $K_{eff}$ , no se aprecia un inicio de tendencia horizontal seguido de una brusca caída del error como se daba en el resto de los casos, esto puede ser porque en este caso el entrenamiento se realiza de forma distinta al no tener que predecir la concentración, sino la constante efectiva. Como se puede ver, alrededor de las 600 épocas el error se estabiliza adoptando una tendencia horizontal.

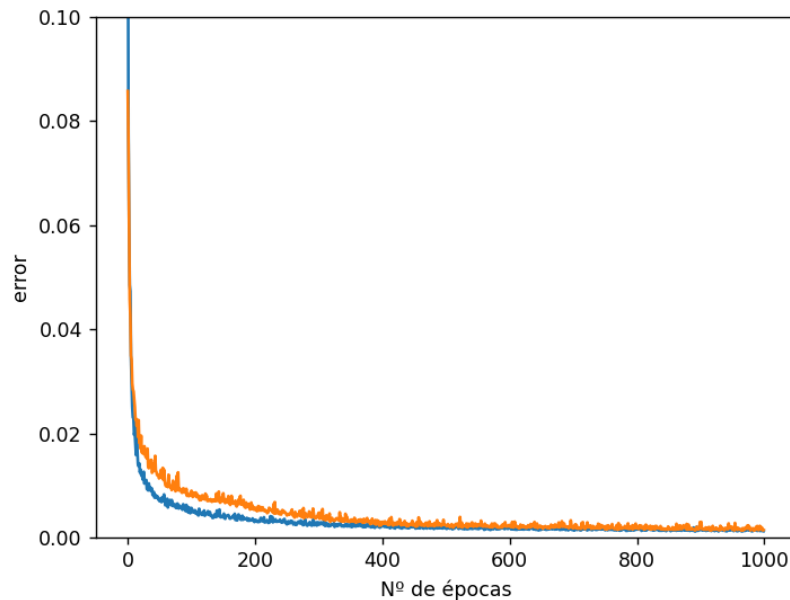


Ilustración 44: Gráfica del error frente al número de épocas

Si se observa la ilustración 45, donde se representan los valores predichos respecto de los valores reales, se puede afirmar que las predicciones son muy buenas, sobre todo si se comparan con el caso inicial. Tal y como en los casos del uranio y el plutonio, la predicción es bastante uniforme, es decir, que funciona de forma adecuada para todos los valores dentro del rango de los respectivos parámetros, no como en el caso del americio, que para valores bajos las predicciones se alejaban un poco.

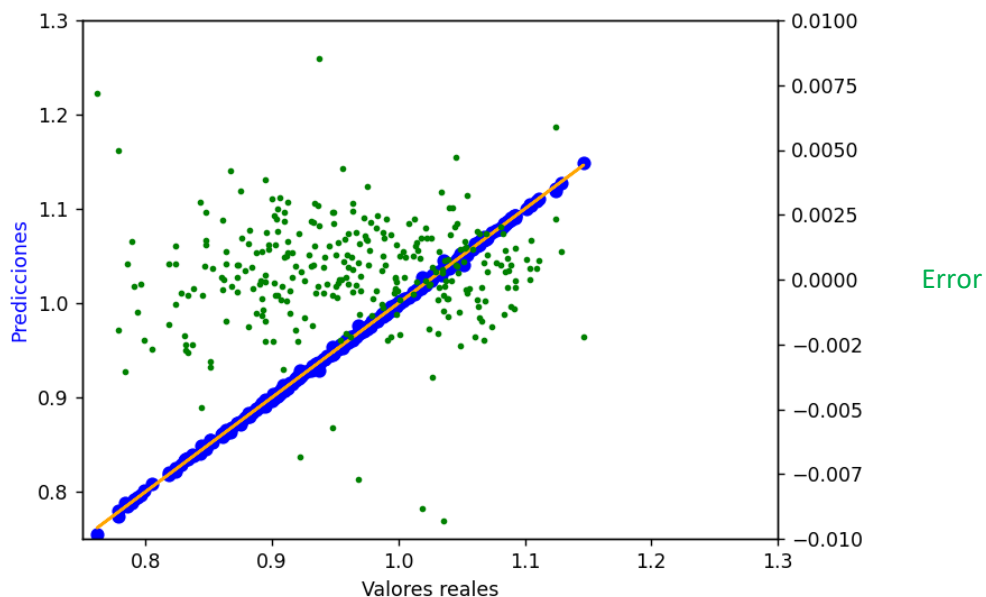


Ilustración 45: Predicciones frente a valores reales de la  $K_{eff}$

Al igual que en los casos anteriores, el aumento del tamaño del set de datos ha llevado consigo una disminución del error, representado en la Tabla 13. por otro lado, como ocurre también con el U-235 y el Pu-239, al ser las predicciones ya buenas de por sí, no se aprecia gráficamente y por eso es necesario centrar más las miradas en los puntos verdes. Para ver esto mejor, la ilustración 46 muestra cómo los errores en las predicciones se van compactando más alrededor del 0.

Tabla 13: Comparación de los errores en los set de datos reducido y ampliado

SET DE DATOS	EMA	EMC
3000 casos	0,0014	3,60E-06
4000 casos	0,001	2,20E-06

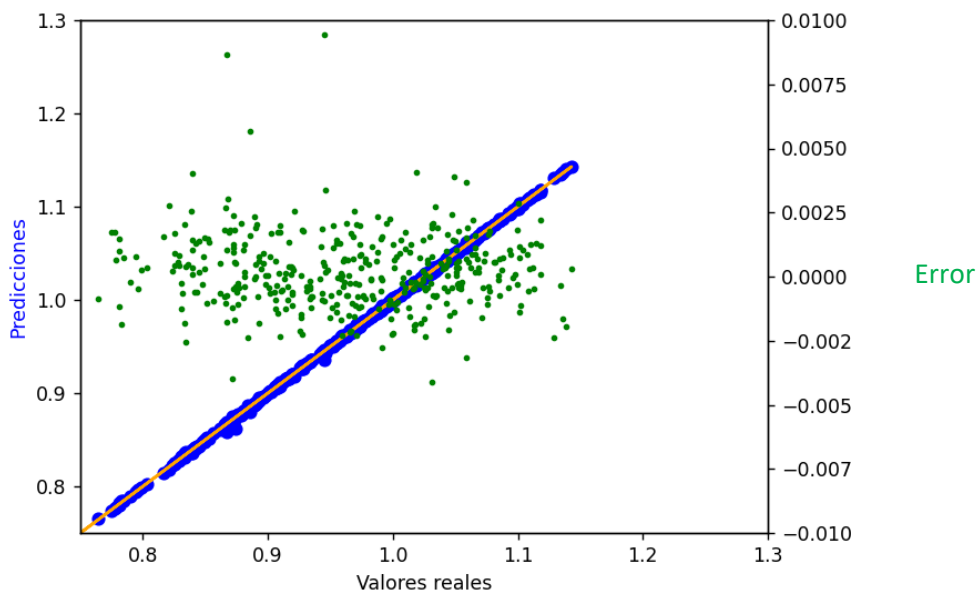


Ilustración 46: Predicciones frente a valores reales y error en la  $K_{eff}$  con el set de datos ampliado

## 9. Planificación del proyecto

Para la realización del proyecto ha sido necesario llevar a cabo una planificación de forma que se establezcan unas pautas para llevarlo a cabo de forma correcta. En este apartado se realizará una breve descripción de las tareas principales y se mostrará el diagrama de Gantt correspondiente.

### 9.1 Descripción de tareas

**P.T.1 Investigación:** Esta tarea comprende la primera fase del proyecto, donde se recoge toda la información necesaria para la realización del proyecto, así como obtener y profundizar en los conocimientos necesarios. Para poder llevar a cabo esta tarea de forma exitosa, es necesario adquirir un conocimiento adecuado de los distintos softwares a utilizar y recopilar información acerca de las redes neuronales para conseguir comprender el concepto y poder ser capaz de desarrollarla.

**P.T.2 Ejecución de casos:** En esta etapa se deben crear los archivos de entrada variando los parámetros necesarios y ejecutar los casos en el programa SCALE. Una vez ejecutados, se extrae de los archivos de salida la información necesaria. Esta ha de disponerse de forma ordenada para facilitar el tratamiento de datos en la siguiente fase.

**P.T.3 Diseño de la red neuronal:** Con los datos necesarios y una vez adquirido el conocimiento necesario para diseñar una red neuronal, los parámetros a tener en cuenta y el tipo de problema que se quiere resolver, se establecerá el set de datos con el cual se llevará a cabo esta fase, el desarrollo y entrenamiento de la red. Dentro de esta fase se aplicará todo lo realizado hasta la fecha para un entrenamiento correcto. Al finalizar el proceso de entrenamiento y prueba de la red, esta estará capacitada para realizar las predicciones necesarias de forma correcta.

**P.T.4 Redacción:** Tras la realización del grueso del proyecto, se tendrá que detallar todo en un documento donde se recogerán los pasos seguidos hasta la finalización del trabajo.

**P.T.5 Revisión:** Revisión por parte del director del TFM con el objetivo de comprobar que el trabajo está bien realizado.

Para comprobar el cumplimiento de la planificación, se han propuesto varios hitos.

**H.1 3000 casos ejecutados**

**H.2 4000 casos ejecutados**

**H.3 Redes diseñadas**

**H.4 Documento redactado**

**H.5 Proyecto finalizado**

## 9.2 Diagrama de Gantt

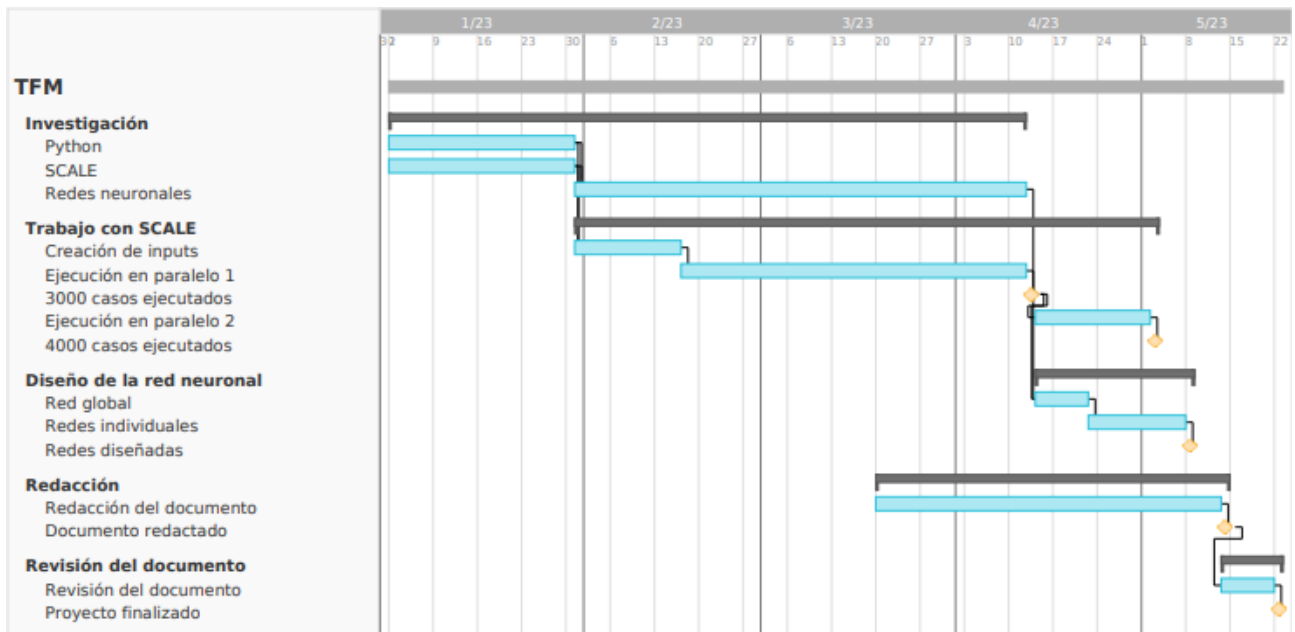


Ilustración 47: Diagrama de Gantt del proyecto

## 10. Descargo de gastos

En este apartado se llevará a cabo el descargo de gastos que ha supuesto la realización de este proyecto. Este se dividirá en horas internas, amortizaciones y gastos.

### 10.1 Horas internas

Las horas internas de este proyecto están conformadas por las horas correspondientes al alumno, las cuales tienen un coste unitario de 30,00 €/h y las horas referentes al director del TFM, cuyo coste asciende a 70,00 €/h.

Tabla 14: Resumen horas internas

Concepto	UD.	Nº UD.	€/UD.	Coste	Coste total
<b>Horas internas</b>					<b>21.500,00€</b>
Director	horas	50	70,00€	3.500,00€	
Alumno	horas	600	30,00€	18.000,00€	

### 10.2 Amortizaciones

Las amortizaciones están divididas entre los equipos utilizados para la realización del proyecto, en este caso solamente el ordenador y la licencia del software necesario.

Tabla 15: Resumen amortizaciones

Concepto	UD.	Nº UD.	€/UD.	Coste	Coste total
<b>Amortizaciones</b>					<b>320,00€</b>
Ordenador	día	80	2,50€	200,00€	
Licencia SCALE	día	60	2,00€	120,00€	

### 10.3 Gastos

Durante la realización de este trabajo, los únicos gastos que se han realizado han sido relativos a material de oficina, como pueden ser una silla, mesa, bolígrafos etc. A pesar de que el trabajo se ha realizado a ordenador, sí se ha hecho uso de cuadernos y bolígrafos para estructurar el trabajo y hacer anotaciones.

Tabla 16: Resumen gastos

Concepto	Coste	Coste total
<b>Gastos</b>		<b>500€</b>
Material de oficina	500€	

## 10.4 Resumen del descargo de gastos

El descargo de gastos viene resumido en la siguiente tabla:

*Tabla 17: Resumen del descargo de gastos*

<b>Concepto</b>	<b>Coste</b>
Horas internas	21.500,00€
Amortizaciones	320,00€
Gastos	500,00€
Subtotal	22.320,00€
Costes indirectos (5%)	1.116,00€
Total	23.436,00€



## 11. Conclusiones

Del estudio realizado anteriormente, se ha comprendido de qué manera se diseña y funciona una red neuronal, además de obtener un conocimiento más profundo sobre el software SCALE y el almacenamiento de combustible gastado.

En primer lugar, por lo descrito en el trabajo ha quedado claro que establecer una sola red neuronal para predecir cada uno de los cuatro parámetros deseados ofrece resultados bastante pobres, sobre todo en términos de la constante de multiplicación efectiva. Un motivo por el cual estos resultados no han sido buenos es porque al ir cambiando los pesos de las neuronas durante el entrenamiento, las características que pueden resultar importantes para la predicción de alguno de los parámetros no sean tan importantes para la predicción de otros y la red no puede utilizar dos pesos distintos en una neurona. Para arreglar este problema, se han diseñado cuatro redes neuronales distintas, una para cada parámetro, lo que ha dado lugar a grandes mejoras en los resultados.

Dentro del diseño de la red neuronal, el utilizar una variación de la función de activación ReLU, denominada PReLU, ha resultado en conseguir menores errores en la mayoría de los casos, siendo la solución apropiada en tres de las cuatro redes desarrolladas.

Las redes neuronales diseñadas ofrecen resultados bastante certeros para las predicciones de los distintos parámetros, sin embargo, el hecho de ampliar el set de datos de 3000 a 4000 casos ha implicado una importante reducción en el error. Esto viene a significar que un aumento del número de casos ejecutados ayuda a la red a entrenar de forma correcta. Durante la realización de este trabajo no se han ejecutado un mayor número de casos por un factor de tiempo, puesto que las simulaciones se realizaban con un ordenador personal con cuatro procesadores, la disponibilidad de un servidor de cálculo habría acelerado la ejecución de los casos y poder alcanzar un tamaño de set de datos importante, el cual hubiera reducido aún más el error.

Otra conclusión que se puede obtener observando los resultados es que no hay una arquitectura de red que sea perfecta, se trata de seguir un procedimiento de prueba y error donde la mejor arquitectura será la que menor error ofrezca.

Como pasos posteriores a este trabajo puede estar el seguir ejecutando casos para poder aumentar el tamaño del set de datos con los que entrenar las distintas redes neuronales y así, ver hasta donde se puede reducir el error. Una vez dicho error sea ya prácticamente irreductible, se podrá diseñar una red neuronal que calcule la  $K_{eff}$  de un contenedor y, mediante el uso de un algoritmo de optimización, optimizar la carga de elementos de combustible irradiados en contenedores para su almacenamiento.

## 12. Bibliografía

(2019). Obtenido de Consejo de Seguridad Nuclear:

[https://www.csn.es/documents/10182/932121/El+combustible+nuclear+y+su+ciclo+\(Monograf%C3%ADa\)](https://www.csn.es/documents/10182/932121/El+combustible+nuclear+y+su+ciclo+(Monograf%C3%ADa))

Brownlee, J. (11 de Agosto de 2020). Obtenido de Machine Learning Mastery:

<https://machinelearningmastery.com/5-step-life-cycle-neural-network-models-keras/>

*Consejo de Seguridad Nuclear*. (s.f.). Obtenido de

<https://www.csn.es/documents/10182/927506/La+energ%C3%ADa+nuclear+%28Monograf%C3%ADa%29>

*Consejo de Seguridad Nuclear*. (s.f.). Obtenido de <https://www.csn.es/almacenamiento-en-piscina>

EDCM. (9 de Noviembre de 2022). El Gobierno descarta finalmente el ATC de Villar de Cañas y hará siete almacenes individuales. *El Español*.

El Gobierno descarta finalmente el ATC de Villar de Cañas y hará siete almacenes individuales. (9 de Noviembre de 2022). *El Español*.

*Foro nuclear*. (diciembre de 2021). Obtenido de <https://www.foronuclear.org/descubre-la-energia-nuclear/energia-nuclear-en-el-mundo/>

*Foro Nuclear*. (2021). Obtenido de <https://www.foronuclear.org/recursos/infografias/generaciones-de-reactores-nucleares/>

*Foro Nuclear*. (2022). Obtenido de <https://www.foronuclear.org/descubre-la-energia-nuclear/como-funciona-una-central-nuclear/>

González, J. M. (17 de Junio de 2020). Obtenido de LinkedIn: <https://es.linkedin.com/pulse/tuneando-los-hiperpar%C3%A1metros-de-una-red-neuronal-1stm-casas-gonzalez>

Hueso, C. (2022). Teoría de Reactores Nucleares.

Liou, J. (03 de diciembre de 2021). *Organismo Internacional de Energía Atómica*. Obtenido de <https://www.iaea.org/es/newscenter/news/que-son-los-reactores-modulares-pequenos-smr#:~:text=Los%20SMR%20permiten%20ahorrar%20costos,la%20demanda%20creciente%20de%20energ%C3%ADa.>

Merino, Á. (28 de Noviembre de 2021). *El Orden Mundial*. Obtenido de <https://elordenmundial.com/mapas-y-graficos/calendario-energia-nuclear-union-europea/?nab=0>

*Ministerio de derechos sociales y agenda 2030*. (2021). Obtenido de <https://www.mdsocialesa2030.gob.es/agenda2030/index.htm>

*Ministerio para la transición ecológica y el reto demográfico*. (2023). Obtenido de <https://energia.gob.es/nuclear/Centrales/Espana/Paginas/CentralesEspana.aspx>

Negm, M. (7 de Marzo de 2019). *Linkedin*. Obtenido de <https://www.linkedin.com/pulse/mlp-vs-cnn-rnn-deep-learning-machine-model-momen-negm#:~:text=So%20MLP%20is%20good%20for,problem%20they%20are%20designed%20for>.

Nuclear, C. d. (2022). Monografía, La energía nuclear.

*Observatorio de Inteligencia Artificial*. (13 de Abril de 2016). Obtenido de <https://observatorio-ia.com/uso-inteligencia-artificial-en-grandes-empresas>

*Office of Nuclear Energy*. (9 de Julio de 2019). Obtenido de <https://www.energy.gov/ne/articles/triso-particles-most-robust-nuclear-fuel-earth#:~:text=What%20is%20TRISO%20Fuel%3F,release%20of%20radioactive%20fission%20products>.

Pinar, C. (09 de 11 de 2022). El gobierno mantiene el calendario de cierre y entierra el proyecto del ATC para que cada central nuclear almacene sus propios residuos. *20 minutos*.

Ponce, J. (10 de Julio de 2021). Obtenido de <https://jahazielponce.com/funciones-de-activacion-y-como-puedes-crear-la-tuya-usando-python-r-y-tensorflow/>

Solans, V., Rochman, D., Brazell, C., Vasiliev, A., Ferroukhi, H., & Pautz, A. (2021). *Optimisation of used fuel canister loading using a neural network and genetic algorithm*. 4: Julio.

Velasco, L. (26 de Abril de 2020). Obtenido de <https://velascoluis.medium.com/optimizadores-en-redes-neuronales-profundas-un-enfoque-pr%C3%A1ctico-819b39a3eb5>

W. A. Wieselquist, R. A. (2020). *SCALE code system*. Abril.