

# Desarrollo de algoritmos basados en inteligencia artificial para clasificación de eventos en los experimentos DEAP-3600

JUAN ANTONIO RAMÍREZ PÉREZ

TRABAJO DE FIN DE MÁSTER

MÁSTER EN FÍSICA NUCLEAR



UNIVERSIDAD DE SEVILLA

Directores:

Miguel Cárdenas-Montes

Vicente Pesudo Fortes

Tutorizado por:

Jose Miguel Arias Carrasco

Curso 2019/2020

## Resumen

La inteligencia artificial y especialmente las redes neuronales están siendo ampliamente utilizadas en física de partículas para extraer información de grandes volúmenes de datos. Por otro lado, uno de los más relevantes desafíos de la física actual es la comprensión de la naturaleza de la Materia Oscura. Uno de los candidatos más prometedores son las partículas masivas débilmente interactivas (WIMP), que se intentan detectar por medio de experimentos basados en argón líquido como DEAP-3600. La gran cantidad de datos generados por parte de estos experimentos tiene que ser procesada y analizada de manera fiable y rápida. Para esta tarea, la inteligencia artificial puede jugar un papel relevante. El objetivo principal de este trabajo es la clasificación, con redes neuronales, de distintos tipos de eventos en el experimento DEAP-3600 en SNOLAB (Canadá). Mediante esta clasificación se pueden separar la señal (eventos producidos por retrocesos nucleares de  $^{40}\text{Ar}$ ) y el fondo (eventos producidos por la desintegración alfa de granos de polvo en suspensión en el detector). Como resultado del trabajo se obtiene una aceptación del 28,1 % para un poder de rechazo del 90 %.

## Abstract

Artificial intelligence, especially neural networks, are being used in particle physics to obtain knowledge from large data volumes. In fact, one of the most relevant challenges in physics today is understanding the nature of Dark Matter. A promising candidates are Weakly Interactive Massive Particles (WIMPs), which are attempted to be detected by liquid argon experiments such as DEAP-3600. The large amount of data generated by these experiments has to be reliably and quickly analyzed. In this task, artificial intelligence could play a relevant role. In this work, the main objective is the classification, with neural networks, of different events in the DEAP-3600 experiment in SNOLAB (Canada). By this classification, the signal ( $^{40}\text{Ar}$  nuclear recoil events) can be separated from the background (alpha decays events produced in dust grains inside the detector). As a result an acceptance of 28,1 %, for a rejection power of 90 %, is obtained.

# Índice

<b>1. Materia Oscura</b>	<b>4</b>
1.1. Evidencias observacionales . . . . .	4
1.1.1. Curvas de rotación de las galaxias . . . . .	4
1.1.2. Lentes gravitacionales . . . . .	5
1.1.3. Anisotropía del CMB . . . . .	6
1.2. Propiedades de la materia oscura . . . . .	7
1.3. Candidatos a Materia Oscura . . . . .	8
1.3.1. Los WIMPs . . . . .	8
1.3.2. Los axiones . . . . .	9
1.3.3. Los neutrinos estériles . . . . .	9
1.3.4. Teoría MOND . . . . .	9
<b>2. Búsqueda directa de materia oscura</b>	<b>10</b>
2.1. Interacción WIMP-Argón . . . . .	11
2.2. Aspectos experimentales . . . . .	13
<b>3. Experimento DEAP-3600</b>	<b>14</b>
3.1. Descripción del detector . . . . .	15
3.2. Rechazo de eventos de fondo . . . . .	16
<b>4. Redes Neuronales</b>	<b>18</b>
4.1. Fundamentos . . . . .	19
4.2. Métricas . . . . .	21
4.3. Fundamentos de las redes neuronales artificiales . . . . .	24
4.4. Fundamentos de las Redes Neuronales Convolucionales . . . . .	25
<b>5. Análisis con Redes Neuronales</b>	<b>27</b>
5.1. Datos desbalanceados . . . . .	27
5.2. Datos balanceados con SMOTE . . . . .	30
5.3. Optimización hiperparámetros . . . . .	33
5.4. Rendimiento para cada muestra de polvo . . . . .	34
5.5. Rendimiento para todas las muestras de polvo juntas . . . . .	39
5.6. Clasificación en función de posición reconstruída . . . . .	39
5.7. Clasificación en función de energía . . . . .	42
5.8. Validación sobre datos de $^{39}\text{Ar}$ . . . . .	43

<b>6. Análisis con Redes Neuronales Convolucionales</b>	<b>45</b>
6.1. Optimización de los hiperparámetros . . . . .	45
6.2. Rendimiento de la CNN . . . . .	45
6.3. Clasificación en función de la posición reconstruída . . . . .	46
6.4. Clasificación en función de la energía . . . . .	48
6.5. Validación sobre datos de $^{39}\text{Ar}$ . . . . .	51
<b>7. Conclusiones y trabajo futuro</b>	<b>52</b>
<b>8. Referencias</b>	<b>54</b>

## 1. Materia Oscura

La materia oscura es una parte fundamental del Modelo Cosmológico Estándar. Según los últimos resultados de la misión PLANCK se observa que la materia oscura representa un  $(26,8 \pm 2,1)\%$  de toda la densidad energética del universo, siendo el 83,7% de toda la materia presente [1]. Su distribución y las condiciones que deben cumplir los potenciales candidatos a materia oscura han sido estudiadas por la comunidad científica en las últimas décadas. Actualmente se conocen varias evidencias experimentales sobre la existencia de la materia oscura, así como algunas de sus propiedades. A pesar de ello, aún no se ha realizado ninguna detección directa de materia oscura [2].

La primera vez que fue introducida la materia oscura fue por el astrónomo suizo Zwicky en 1933, cuando estudiaba las velocidades de dispersión de ocho galaxias del cúmulo de Coma [3,4]. Aplicando el teorema del virial estimó la masa del cúmulo, obteniendo un valor unas 400 veces por debajo del esperado. Su conclusión fue que debía existir una “materia oscura” que no fuese observable y que no interaccionase con la radiación electromagnética.

### 1.1. Evidencias observacionales

Existe una larga lista de evidencias observacionales que aseguran la existencia de materia oscura. A continuación se realizará un repaso sobre las más significativas.

#### 1.1.1. Curvas de rotación de las galaxias

La velocidad de un cuerpo (estrella o gas) en órbita de una galaxia es menor a medida que se aleja del centro de esta [5]. Si se representa la curva de rotación de las galaxias no existe un decrecimiento de la velocidad a mayor radio, sino que tiende a un cierto valor constante. Este fenómeno fue descubierto por Rubin y Ford en 1978 (Figura 1). Observaciones similares de curvas de rotación planas se han encontrado en muchas de las galaxias estudiadas, incluyendo la Vía Láctea.

Para poder explicar este fenómeno, las galaxias deben contener mucha más masa de la que puede ser explicada por el brillo de objetos estelares residentes en los discos galácticos. Este déficit persiste al añadir el polvo o gas interestelar, cuya presencia es observable en el espectro de rayos X. Las galaxias deben poseer halos con gran cantidad de materia oscura distribuida. Este perfil de velocidad se representa en la Figura 1, donde la materia bariónica sola no es capaz de explicar la curva de rotación galáctica. Sin embargo, agregar materia oscura en el halo galáctico permite un buen ajuste de los

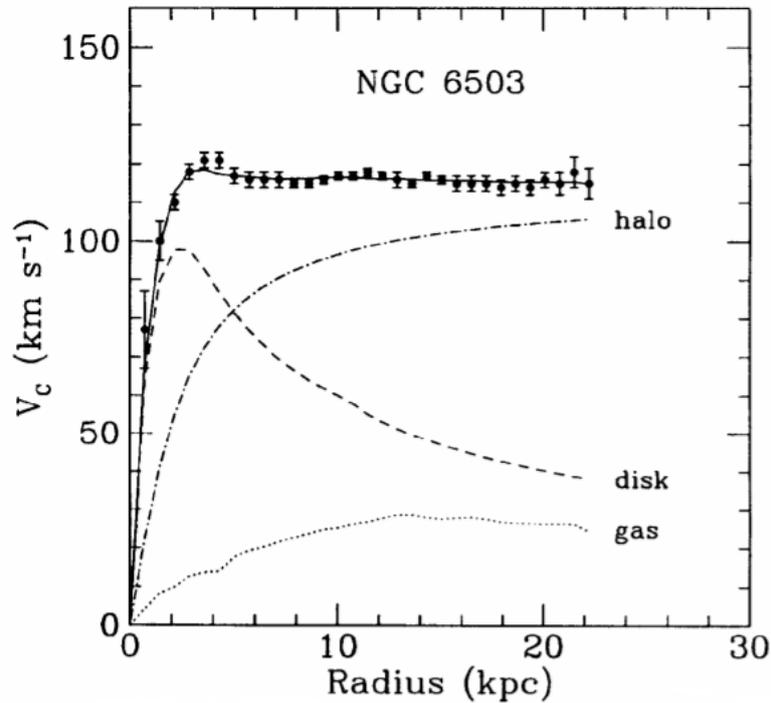


Figura 1: Curva de rotación galáctica para la galaxia NGC-6503. Se observan las contribuciones de disco y gas más la contribución del halo de materia oscura necesaria para los datos [5].

datos experimentales.

### 1.1.2. Lentes gravitacionales

Las lentes gravitacionales son un efecto que surge cuando un objeto masivo se encuentra en la línea de visión entre el observador, que en este caso es la Tierra, y un objeto ligero [6]. Debido al campo gravitatorio generado por el objeto masivo (objeto 1), los rayos de luz serán desviados de su trayectoria, produciendo una deformación de la imagen observada del objeto 2. Este efecto puede ser observado en la Figura 2.

A través del grado de deformación de la imagen se pueden determinar distintas características del objeto cuya gravedad produce la desviación de la luz, como por ejemplo, su masa. A través de varios experimentos se obtuvieron masas mayores a las que corresponderían si solo existiese materia bariónica o luminosa. Se deduce así la existencia de la materia oscura.

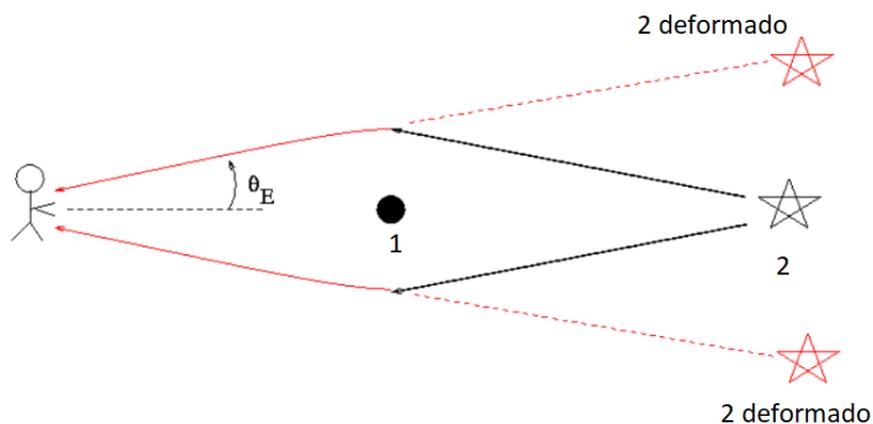


Figura 2: Geometría básica de una lente gravitacional. Un rayo de luz proveniente del objeto 2 es deflektado por el objeto 1. En el caso de las lentes gravitacionales fuertes se produce un anillo alrededor del objeto 1 conocido como “anillo de Einstein” [6].

### 1.1.3. Anisotropía del CMB

La *radiación de fondo* o fondo cósmico de microondas (CMB) nos muestra el Universo unos 380.000 años después del Big Bang, cuando la temperatura había descendido hasta 3.000 °C. [2]. Este evento se conoce con el nombre de *recombinación*. En esta época el Universo se volvió transparente. A energías lo suficientemente altas, se reionizaba constantemente el hidrógeno. Al reducirse la energía de los electrones, la tasa de ionización disminuyó, permitiendo a los fotones viajar libremente.

Debido a la expansión del universo, los fotones de aquella época son hoy menos energéticos. Hoy corresponden con el espectro de un cuerpo negro de 2,725 K. La radiación de fondo no es perfectamente isotrópica, existen fluctuaciones del orden de las diezmilésimas de grado.

Se pueden predecir las fluctuaciones de la temperatura en función de la distancia angular, como podemos observar en la Figura 3. En ella, a través de la altura relativa y de la distancia entre máximos se puede determinar la edad del Universo y su tasa de expansión. También es posible obtener los parámetros de curvatura y de densidades de la materia bariónica, materia oscura y energía oscura [1]. De aquí se puede extraer que la materia oscura representa un 26,8 %, mientras que la materia bariónica es tan solo un 4,9 %.

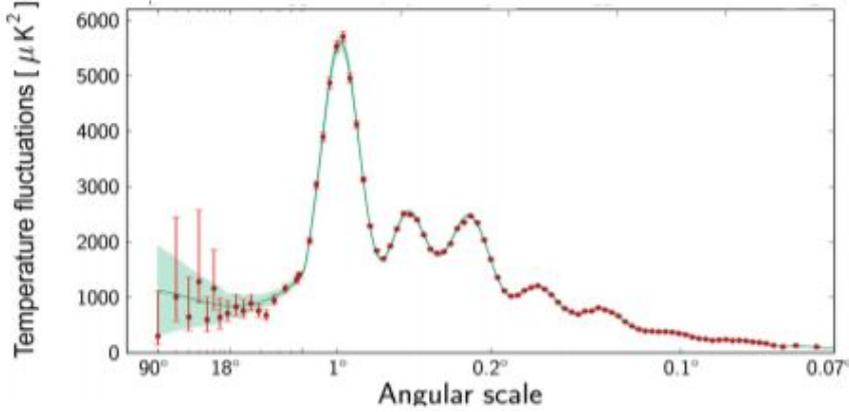


Figura 3: Fluctuaciones de la temperatura del CMB [7].

## 1.2. Propiedades de la materia oscura

A pesar del desconocimiento de la naturaleza de la materia oscura, las evidencias observacionales permiten conocer sus propiedades [2]. Las principales propiedades son:

- La materia oscura interactúa muy débilmente con la materia bariónica. Sólo se la ha podido observar interactuar gravitacionalmente hasta ahora. Si existiese otro tipo de interacción, tendría que ser muy débil o ya se habría detectado la materia oscura de forma directa. También debe interactuar muy débilmente consigo misma, afirmación que podemos extraer a través del cúmulo de la Bala, fruto de la colisión de dos cúmulos de galaxias [8]. Se ha podido observar a través de lentes gravitacionales que en este cúmulo las componentes de materia oscura han atravesado la colisión sin interactuar entre sí, mientras que el gas intergaláctico ha sido frenado.
- La materia oscura no puede estar formada por partículas ordinarias como protones, neutrones y electrones, o los cálculos sobre la nucleosíntesis primitiva serían contradictorios respecto a los datos observacionales. En el caso que la materia oscura estuviese formada por materia bariónica, el universo temprano habría sido más denso. La conversión de materia en  ${}^4\text{He}$  habría sido más eficiente. Tendría que existir menos deuterio del que se observa.
- La materia oscura debe ser eléctricamente neutra. Si la materia oscura estuviese formada por partículas cargadas eléctricamente, esta tendría que haber formado parte del plasma primitivo. A través del espectro de fluctuaciones de temperatu-

ra del CMB, se puede extraer que la materia oscura no formó parte del plasma primordial de la materia ordinaria.

- Los porcentajes de materia oscura que pueden inferirse de la época de la recombinación y la época actual son compatibles, por lo que la materia oscura debe ser estable. Así que la vida de la materia oscura debe ser como mínimo 13.800 millones de años, la edad del universo.
- La materia oscura debe ser fría, con una velocidad de movimiento lo suficientemente pequeña para que hubiese podido agruparse por efecto de su propia gravedad.

### 1.3. Candidatos a Materia Oscura

Existe una gran variedad de candidatos a materia oscura. A continuación se analizan las distintas posibilidades, así como alguna hipótesis alternativa a la existencia propia de la materia oscura.

#### 1.3.1. Los WIMPs

Los WIMPs (*Weakly Interacting Massive Particle*), son partículas masivas, con una masa entre los 2 GeV y los 100 TeV. Estas partículas sólo presentarían interacción débil. Los WIMPs estarían de acuerdo con los cálculos teóricos sobre la abundancia de materia oscura en el universo.

Para obtener la abundancia de materia oscura a través de la producción térmica se requiere una sección transversal  $\langle \sigma v \rangle \approx 3 \cdot 10^{-26} \text{ cm}^3 \text{ s}^{-1}$ . Este valor es el que se espera para una partícula que interactúa débilmente, con un rango de masas como el dicho anteriormente. Existen predicciones que aseguran una nueva partícula con estas propiedades. Esta coincidencia se conoce como *milagro WIMP*.

Los WIMPs son los candidatos más estudiados hasta la fecha para la materia oscura, tanto desde el punto de vista teórico como el experimental [9]. Dentro del Modelo Estándar no existen candidatos posibles para ser un WIMP. Más allá del Modelo Estándar existen candidatos como las partículas supersimétricas. La partícula supersimétrica más ligera, el neutralino, posee las características perfectas de un WIMP: es estable, neutro, solo interactúa débilmente, su masa está en el rango adecuado y da lugar a materia oscura fría.

Otros modelos teóricos más allá del Modelo Estándar que también dan lugar a WIMPs son los modelos con dimensiones espaciales extras, los modelos de Higgs compuestos o los modelos de Higgs múltiples [2].

### 1.3.2. Los axiones

Los axiones son partículas propuestas para resolver el problema teórico de la simetría de paridad y carga de las interacciones fuertes [10]. Estas partículas se forman mayoritariamente por mecanismos no térmicos, lo que hace que sean de naturaleza no relativista a pesar de ser muy ligeros. Son partículas que se ajustan perfectamente a las propiedades de la materia oscura, siendo partículas muy estables, de masa muy pequeña ( $10^{-7}$  eV) y con una interacción extremadamente débil [2].

En caso de existir, el axión debería transformarse, en el seno de un campo electromagnético, en un fotón, y viceversa. Esta propiedad es explotada en los experimentos de búsqueda directa de axiones (IAXO) [11].

### 1.3.3. Los neutrinos estériles

Los neutrinos estériles son partículas candidatas a ser materia oscura que fueron postuladas en 1993 por Dodelson y Widrow. Estas partículas no se encontrarían en los primeros instantes del Universo en equilibrio térmico [12]. Los neutrinos estériles poseen interacción fuerte y débil, además de carga neutra. También poseen masa distinta de cero. Son partículas muy difíciles de detectar.

Los neutrinos estériles pueden generarse en un amplio rango de temperaturas y un amplio rango de masas, que van desde unos pocos keV (materia oscura caliente) hasta valores mucho mayores (materia oscura fría) [13]. En experimentos de energías mayores que su masa, los neutrinos participarían en todos los procesos donde lo harían los neutrinos ordinarios. Un experimento dedicado a la búsqueda de neutrinos estériles es LSND [14].

### 1.3.4. Teoría MOND

La teoría de dinámica newtoniana modificada (teoría MOND), fue propuesta por Milgrom en 1983 [15]. Se trata de una modificación de la teoría gravitacional de Newton, donde en el límite de aceleraciones muy pequeñas se podría explicar la curva de rotación de las galaxias sin necesidad de postular la existencia de la materia oscura.

Según esta teoría la segunda ley de Newton debía ser sustituida por la siguiente expresión:

$$F = m\mu\left(\frac{a}{a_0}\right) a \quad , \quad (1)$$

donde  $a_0$  sería una nueva constante física que debería ser obtenida experimentalmente y  $\mu(x)$  una función cuyos valores asintóticos son:

$$\mu\left(\frac{a}{a_0}\right) = \begin{cases} 1 & |a| \gg a_0 \\ \frac{a}{a_0} & |a| \approx a_0 \end{cases}, \quad (2)$$

Sin embargo, esta teoría no es capaz de explicar todas las evidencias observacionales, como el cúmulo de la Bala, las lentes gravitacionales o las anisotropías del CMB.

## 2. Búsqueda directa de materia oscura

La búsqueda de materia oscura se puede clasificar en tres tipos de experimentos: búsqueda directa, búsqueda indirecta y búsqueda en aceleradores (Figura 4).

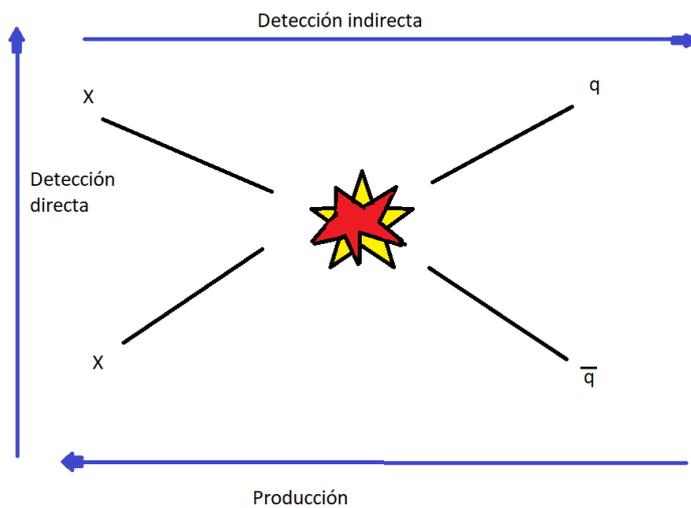


Figura 4: Estrategias de búsqueda de materia oscura. Según el sentido en el que se lea el diagrama se obtiene un método u otro.

- En la búsqueda indirecta se espera detectar una aniquilación entre dos WIMPs en el universo, observando un exceso de flujo de partículas del Modelo Estándar. Algunos de estos experimentos son AMS-02 [16] o DAMPE [17].
- En la búsqueda en aceleradores, se producen colisiones a altas energías entre partículas del Modelo Estándar. Se estudian los déficits en la energía reconstruida tras la colisión. En este aspecto se realizan búsquedas en grandes aceleradores, como en el LHC [18]. A fecha de hoy no se ha encontrado ningún evento significativo.

- En el tercer y último método de búsqueda, la búsqueda directa, se pretende detectar el paso de un WIMP cuando atraviesa la Tierra y produce una dispersión con los núcleos de un material que hace de blanco. Existen diversas técnicas, como el empleo de cristales (ANAIS, CDMS), cámaras de burbujas (PICASSO, SIMPLE) o detectores de gases nobles en estado líquido como ArDM [19], XENON-1T [20], DarkSide-50 [21] o DEAP-3600 [22].

A la hora de realizar una búsqueda directa de materia oscura, se suelen emplear gases nobles como el argón o el xenón. Estos elementos poseen excelentes propiedades como centelleadores, por lo que son transparentes a su propia luz de centelleo. También poseen una gran capacidad de recolección de carga de ionización y permiten una buena resolución de la reconstrucción de la posición. Al ser líquidos son escalables a grandes detectores, por lo que permiten el uso de técnicas de discriminación de partículas. Estos experimentos se realizan en laboratorios subterráneos para minimizar el flujo de rayos cósmicos. La reducción en el flujo de muones por los cientos de metros de roca hacen posible el uso de esta tecnología para la búsqueda de materia oscura.

En particular, en los detectores de argón es posible desarrollar la técnica de discriminación en forma de pulso (PSD) para la identificación de partículas y rechazo del fondo. Los depósitos de energía en argón pueden producir dímeros de argón en estado singlete o triplete. Los retrocesos nucleares, producidos por neutrones y que se espera que produzcan los WIMPs, excitan preferentemente estados singlete. Por otro lado, los retrocesos electrónicos, producidos por radiación beta y gamma, excitan preferentemente estados triplete. Teniendo en cuenta las diferencias de tiempos de decaimiento entre ambos estados se evalúa la forma del pulso. Se distingue así entre un electron recoil (background) o un nuclear recoil (señal de interés).

### 2.1. Interacción WIMP-Argón

En los experimentos de búsqueda directa de materia oscura, la interacción entre el WIMP y un núcleo del elemento que se use como detector es un choque elástico (Figura 5). Tras el choque, el núcleo pasará de estar inicialmente en reposo a adquirir una energía de retroceso  $E_R$  [23]. Esta energía de retroceso dependerá de la masa de las dos partículas y del ángulo de *scattering*  $\theta$ :

$$E_R = \frac{4M_W M_N}{(M_N + M_W)^2} \frac{1 - \cos\theta}{2} E_0 \quad , \quad (3)$$

donde  $M_W$  y  $E_0$  son la masa y energía cinética del WIMP, el cual debido a su masa

es considerado no relativista, y  $M_N$  la masa del núcleo. En el experimento DEAP-3600 (*Dark Matter Experiment using Argon Pulseshape Discrimination*) los núcleos son de argón.

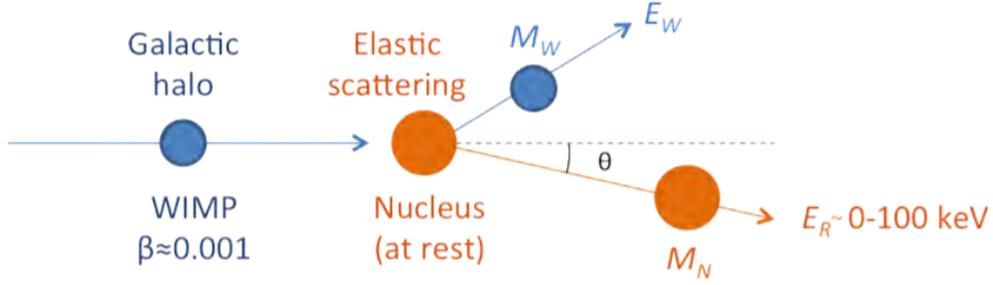


Figura 5: Dispersión elástica de un WIMP con un núcleo [19].

En estos modelos, la dispersión elástica de WIMPs de masas entre 0.1 - 1 TeV con núcleos produce retrocesos nucleares de baja energía, inferior a los 100 keV [22]. Estos sucesos son de baja energía, por lo que es posible que no sean detectados debido al fondo que posee el detector.

En el límite no relativista, hay dos posibilidades de canales de interacción WIMP-núcleo: una dependiente del spin del núcleo y otra independiente del spin [24]. Debido a que el argón posee los dos números cuánticos pares ( $A=18$ ,  $Z=40$ ), su spin será nulo, por lo que sólo podremos obtener una contribución independiente del spin en un detector de argón. La sección eficaz para el canal independiente del spin (SI) se expresa como:

$$\sigma_0^{SI} = \frac{4\mu^2}{\pi} \{Zf_p + (A - Z)f_n\} \quad , \quad (4)$$

donde  $\mu$  es la masa reducida del sistema WIMP-blanco,  $Z$  es el número atómico y  $A$  es el número másico.  $f_n$  y  $f_p$  son las funciones de acoplamiento de los WIMPs a los neutrones y a los protones [25]. Se observa que la sección eficaz es linealmente dependiente con los números de protones y de neutrones, por lo que a mayor tamaño del núcleo, esta será mayor.

En la Figura 6 se muestran distintos resultados a través de diferentes experimentos de materia oscura. Estas curvas de exclusión representan, con un límite de confianza del 90 % las secciones eficaces de interacción con la materia oscura rechazadas para cada masa por cada experimento. Para el experimento DEAP-3600 los últimos resultados imponen un límite superior de  $1,5 \cdot 10^{-44} \text{ cm}^2$  para un WIMP cuya masa sea de 100

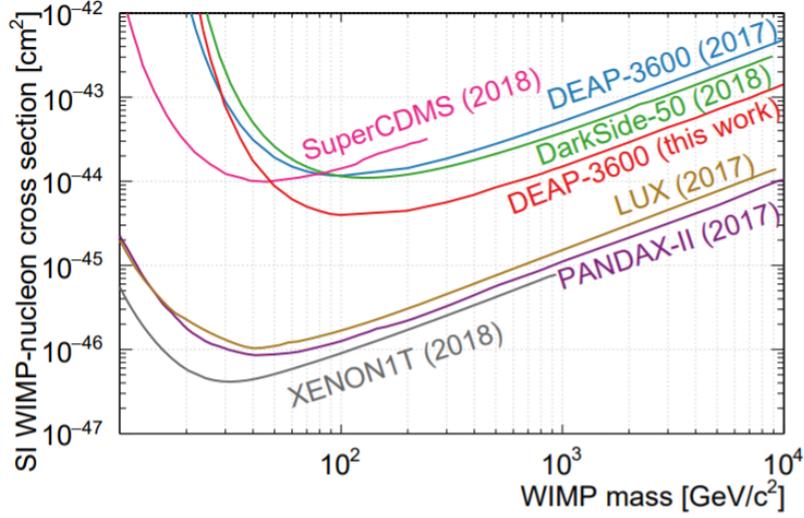


Figura 6: Curvas de exclusión en las secciones eficaces independientes del spin WIMP-nucleón obtenidos a través de varios experimentos [17].

GeV [17]. En el caso de LUX, XENON-1T y PANDAX-II han finalizado su fase de toma de datos, mientras que en DEAP está en curso su análisis y toma de datos. Se esperan obtener límites del mismo orden en un futuro cercano.

## 2.2. Aspectos experimentales

Tras la dispersión elástica entre el WIMP y el núcleo de argón, este último adquiere una energía junto a un movimiento [19]. Este movimiento puede excitar o ionizar otros átomos de argón cercanos, llegando a producir estados moleculares excitados (dímeros), que decaen mediante la emisión de fotones de 9,8 eV. Estos fotones no poseen la energía suficiente para excitar un átomo de argón a su primer estado excitado (15,76 eV), por lo que el argón es transparente a su propia luz de centelleo.

Cuando un rayo cósmico impacta sobre un núcleo de  $^{40}\text{Ar}$ , es posible que se produzca un isótopo radiactivo, el  $^{39}\text{Ar}$ , produciendo una contaminación natural en el argón atmosférico. Este isótopo produce decaimientos  $\beta$  con una actividad de  $(0,95 \pm 0,05)$  Bq/kg [22]. En los decaimientos se emiten electrones, que generarán retrocesos electrónicos con los núcleos de  $^{40}\text{Ar}$  al ser decaimientos  $\beta$ .

Para distinguir los retrocesos nucleares generados por la dispersión WIMP-núcleo de los retrocesos electrónicos se emplea la técnica de discriminación por forma de pulso (PSD). Esta técnica consiste en medir el número de fotoelectrones rápidos y tardíos. Para ello

se define el parámetro  $F_{prompt}$ , el cuál es la fracción de carga de un evento registrada en una pequeña ventana de tiempo sobre el tiempo total del evento. Esta ventana se sitúa al principio del evento. Su duración es de 60 ns, mientras que el pulso total dura 10  $\mu$ s. Los retrocesos electrónicos poseerán un valor por debajo de 0,3 mientras que los retrocesos nucleares suelen tener un valor por encima de 0,6 (Figura 7).

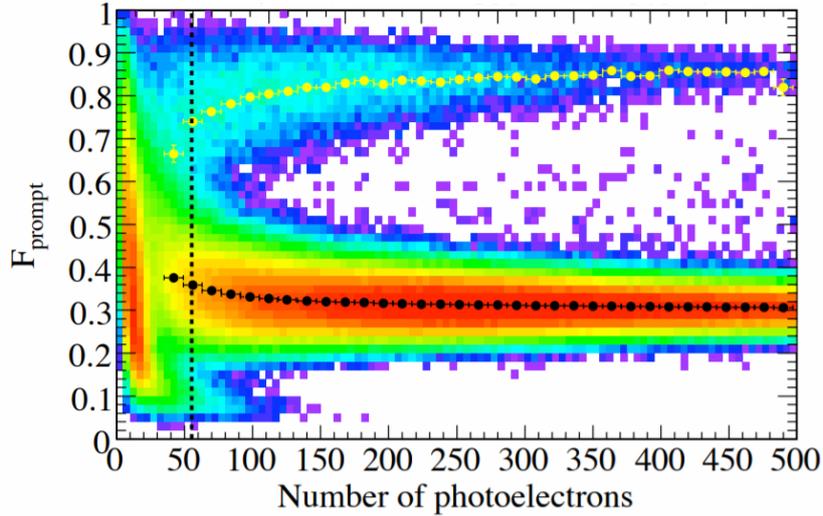


Figura 7: Ejemplo visual de distinción entre ambos tipos de retrocesos mediante el parámetro  $F_{prompt}$ . Los retrocesos nucleares poseen valores en torno a 0,8 mientras que los retrocesos electrónicos alrededor de 0,3 [26].

En los detectores de argón existen otras fuentes de fondo, las cuáles incluyen a rayos  $\gamma$ , neutrones, partículas  $\alpha$  o partículas  $\beta$ . Las alfa, beta y gamma provienen de las cadenas de desintegración del  $^{238}\text{U}$  y  $^{232}\text{Th}$  de los materiales del detector. Las alfas al interactuar con los materiales pueden producir neutrones. Este fondo es menos abundante pero más peligroso porque no se puede descartar con técnicas de PSD. También puede existir contaminación interna, pudiendo incluso existir partículas de polvo en el volumen activo del detector. Estas contribuciones al fondo producen señales que pueden confundirse con las esperadas para los retrocesos nucleares y son las que acaban limitando la sensibilidad de los detectores de búsqueda de materia oscura. Las redes neuronales han demostrado ser una herramienta extremadamente útil para mejorar la ratio señal-fondo [27].

### 3. Experimento DEAP-3600

El experimento DEAP-3600 es un experimento de detección directa de materia oscura con argón líquido como volumen activo [19]. El detector opera a 2km bajo una mina

excavada para la extracción de níquel en el laboratorio de SNOLAB (Sudbury, Ontario, Canadá), otorgándole así la protección de unos 6000 m.w.e frente a los rayos cósmicos (metros equivalentes de agua).

### 3.1. Descripción del detector

Se puede observar un corte transversal del detector en la Figura 8. Este detector consiste en una vasija esférica de acrílico (PMMA) de 5 cm de grosor y un diámetro interno de 1,7 metros [22]. En su interior se pueden albergar hasta 3600 kg de argón, aunque sólo se albergan  $(3279 \pm 96)$  kg. A 55 cm sobre el ecuador de la esfera, existe una interfase líquido-gas. Los 30 cm superiores de la esfera acrílica están llenos con argón gaseoso.

Alrededor de la esfera acrílica existen 255 fotomultiplicadores (PMTs), los cuales registran la luz emitida por los eventos en su interior. Estos están separados de la esfera 45 cm por guías de luz, las cuales transportan los fotones visibles desde la esfera acrílica hasta los PMTs. Esta separación es debida a que los PMTs trabajan a temperaturas próximas a la ambiental, mientras que el argón debe estar a temperaturas criogénicas. La separación también ayuda a apartar los PMTs, que suelen ser sucios, del volumen de detección, reduciendo así el fondo. Para mantener las temperaturas criogénicas del argón, está instalado en la parte superior del detector un cuello donde un serpentín relleno de nitrógeno líquido condensa argón líquido. Entre la vasija de acrílico y la carcasa de acero esférica existen capas alternas de polietileno y poliestireno que sirven como blindaje para los neutrones.

La superficie interna de la esfera acrílica está recubierta con una capa de 3 micras de 1,1,4,4-tetraphenyl-1,3-butadiene o TPB ( $C_{28}H_{22}$ ) para convertir la radiación ultravioleta (128 nm) de centelleo en una luz azul en torno a 420 nm [28]. Esta luz es visible para los PMTs, ya que estos no son sensibles a la radiación de centelleo del argón de 128 nm.

El argón en estado gaseoso posee distintos mecanismos de centelleo, además que su masa es despreciable comparada con el argón líquido, por lo que se simplifica el análisis. Se deja al gas fuera de la región fiducial para la búsqueda de WIMPs. El volumen fiducial es definido por los eventos reconstruidos por debajo de la interfase, es decir, con  $Z_{rec} < 55$  cm, y dentro de un radio de 63 cm, es decir,  $R_{rec} < 63$  cm.

Toda la infraestructura está contenida en un blindaje de acero inoxidable y después está sumergido en un tanque de 1490 m<sup>3</sup> de agua. El blindaje de acero supone un soporte

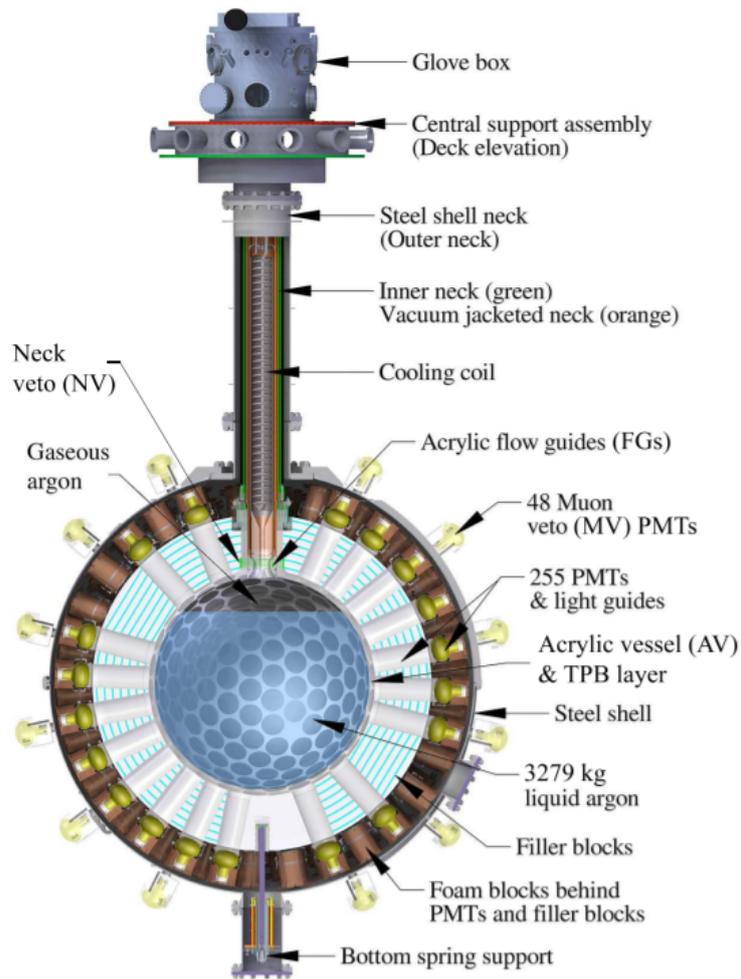


Figura 8: Sección transversal del detector DEAP-3600, donde se observan las distintas partes que lo componen [22].

para 48 PMTs que apuntan hacia afuera y se usan para detectar en el tanque de agua la radiación de Cherenkov producida por los muones de alta energía. Actúan como un veto activo. Además, el tanque de agua contribuye a reducir el flujo de rayos  $\gamma$  y de neutrones procedentes del exterior.

### 3.2. Rechazo de eventos de fondo

Una de las contribuciones al fondo más difíciles de discriminar son los eventos generados por partículas  $\alpha$  provenientes del interior del cuello del detector [22]. DEAP-3600 ha demostrado una excelente capacidad para discriminar fondos beta y gamma a través de PSD y de las radiaciones externas en la superficie del detector gracias a una precisa

reconstrucción de posición, acompañada con una fiducialización del volumen activo. Al eliminar las fuentes más evidentes de radiación, se han hecho evidentes otras que a priori hubiesen sido subdominantes.

Las superficies del cuello del detector poseen isótopos radiactivos provenientes de la cadena del radón como el  $^{210}\text{Po}$ . Este isótopo emite partículas  $\alpha$  que producirán luz de centelleo al contactar con el argón condensado que rodea las superficies de las guías del cuello, como se observa en la Figura 9. Estos eventos producen un patrón de luz característico, aunque no exactamente igual en todos los casos.

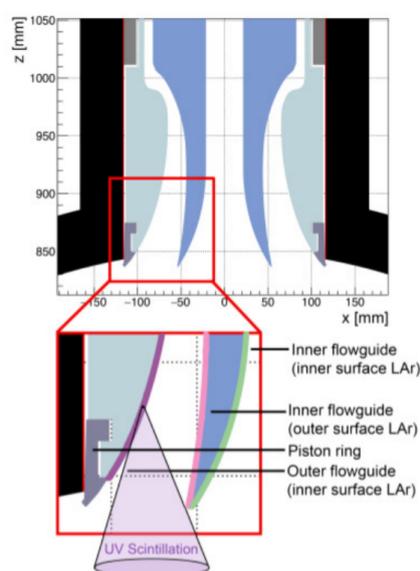


Figura 9: Interior del cuello del detector donde se producen los decaimientos  $\alpha$  [22].

El patrón de luz es difícil de capturar con técnicas de análisis tradicional, pero es identificable topológicamente mediante redes neuronales y otras técnicas de *machine learning* como *Random Forest* y *Boosted Decision Trees*. Estos algoritmos permiten aceptancias del 50% para eventos simulados de WIMPs con un rechazo del fondo del 99.9%. Mediante técnicas tradicionales se introducen 10 veces más eventos de fondo en la región de interés.

Tras esta importante reducción, la presencia de partículas de polvo en suspensión en el argón líquido se ha convertido en el fondo más problemático.

Mediante el uso de redes neuronales, la contribución al fondo asociada a los eventos de cuello ha sido solucionada relativamente bien. Por este motivo, la presencia de muestras de polvo en el interior del volumen fiducial del detector se ha convertido en el fondo

más problemático. La hipótesis más plausible es que sean pequeñas partículas del metal de las botellas que almacenan el nitrógeno (Norita). Recientemente se demostró que los filtros que se utilizaron pueden dejar pasar partículas de polvo de radios inferiores a 25 micras.

Debido al contenido de material radioactivo en las muestras de polvo, se producen partículas  $\alpha$ , y a su vez luz de centelleo cuando depositan su energía con el argón líquido del detector. La luz de centelleo generada la partícula alfa que emite el grano de polvo se propaga isotrópicamente. Como la interacción sucede en la inmediatez del grano, él mismo produce una sombra. La presencia de esta sombra es una característica que podría ser explotada por una red neuronal para diferenciar eventos de retrocesos nucleares y eventos asociados a granos de polvo.

Cuando la luz de centelleo producida por las partículas  $\alpha$  del fondo de partículas de polvo alcanza el detector, el patrón temporal de la señal producida será muy similar a la esperada para los retrocesos nucleares. Debido a esta gran similitud es necesario aplicar un algoritmo que sea capaz de rechazar los eventos de fondo, manteniendo una alta selección de señales buenas. En este trabajo nos centramos en desarrollar una red neuronal que rechace los eventos de fondo de las partículas de polvo, debido a sus buenos resultados reduciendo el fondo de eventos de cuello. Gracias a ello, la curva de exclusión del experimento DEAP-3600 será más sensible, reduciendo así el fondo.

Las partículas de polvo que se encuentran en el detector poseen distintos tamaños, entre 1 y 25 micras. También tienen un movimiento browniano debido a la agitación térmica del argón líquido, lo que provoca que posean una distribución uniforme a lo largo del volumen fiducial del detector. También es necesario tener en cuenta que la partícula alfa puede salir en cualquier dirección. Sin embargo, los eventos de cuello forman un cono de luz hacia la parte inferior del detector. Esta es una diferencia clave entre ambas contribuciones de fondo.

## 4. Redes Neuronales

Las redes neuronales son uno de los algoritmos más populares de aprendizaje automático, debido a que han demostrado excelentes resultados en campos tan amplios como la medicina o las finanzas [29]. En este apartado se realiza un estudio sobre las redes neuronales artificiales y las redes neuronales convolucionales. Una red neuronal artificial es el conjunto de varias unidades de cómputo denominadas neuronas en distintas capas, ya que sus estructuras están inspiradas en las neuronas humanas. Una red neuronal convo-

lucional posee aprendizaje supervisado que procesa sus capas imitando al cortex visual del ojo humano para identificar distintas características en las entradas.

#### 4.1. Fundamentos

Una neurona es una unidad de cómputo que evalúa una función de activación  $h(z)$  sobre el producto escalar entre un vector de entrada por un vector de pesos  $\vec{\theta}$  más un término independiente  $\theta^0$ , la ordenada en el origen [30]. El valor de salida de una neurona será por tanto:

$$a = h \left( \theta^0 + \sum_{i=1}^m \theta^i x^i \right) \quad , \quad (5)$$

donde  $i$  es el índice de las variables de entrada,  $m$  el número de estas y  $x^i$  el vector de entrada. La entrada de cada neurona de la capa  $c$  será por tanto la salida de cada neurona la capa anterior,  $c - 1$ . Con esto en cuenta podemos reescribir la ecuación (5) :

$$a_c^j = h \left( \theta_c^{j,0} + \sum_{i=1}^m \theta_c^{j,i} a_{c-1}^i \right) \quad , \quad (6)$$

donde  $j$  es el índice de la neurona de la capa  $c$ .

La primera capa se conoce como capa de entrada o capa de lectura, donde el número de neuronas es el mismo que el número de variables que representan los datos. La última capa se conoce como capa de salida, donde se obtiene el resultado final de la red neuronal. Las distintas capas intermedias se conocen como capas ocultas. En una red neuronal completamente conectada como la que usamos, todas las neuronas de una capa se conectan con todas las neuronas de la capa siguiente. Las distintas topologías de las redes neuronales completamente conectadas quedan definidas por el número de neuronas por capa y por el número de capas [31]. En la Figura 10 se puede observar un esquema de red neuronal.

La función de activación calcula el estado de actividad de una neurona, transformando la entrada global en un valor (estado) de activación. En la capa de salida, la función de activación depende del problema que quiera solucionar. En problemas de regresión numérica se suele emplear una función lineal, mientras que en problemas de clasificaciones categóricas se emplea la función sigmoide, con un rango comprendido normalmente entre 0 y 1. Esto es una buena representación de que la neurona puede estar totalmente

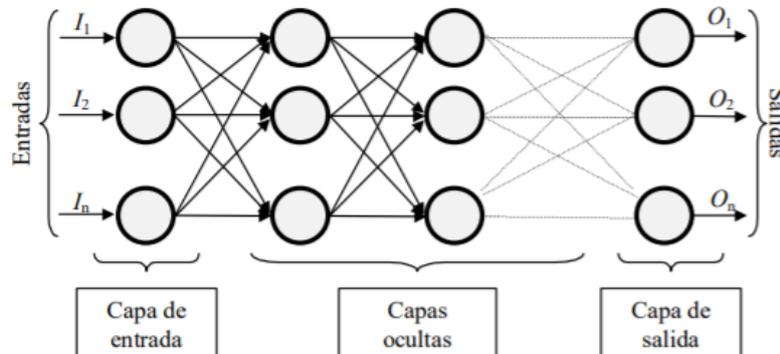


Figura 10: Ejemplo de una red neuronal totalmente conectada [27].

inactiva (0) o totalmente activa (1). Se tienen ventajas para el tratamiento matemático respecto la función escalón, al ser una función regular.

El objetivo del aprendizaje consiste en encontrar la configuración de los pesos que corresponde al mínimo global de la función error o de coste,  $J$ . Esta función variará según el problema a resolver, en los problemas de regresión numérica se puede emplear el error cuadrático medio, mientras que en los problemas de clasificación binaria se suele emplear la entropía cruzada, cuya expresión es:

$$J = \frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad , \quad (7)$$

donde  $y_i$ ,  $\hat{y}_i$  son los valores reales y predichos por la red neuronal para cada ejemplo  $i$  y  $N$  el número total de ejemplos que se han empleado para entrenar la red neuronal.

El método de optimización de los pesos de la red  $\theta$  es el denominado *descenso por el gradiente*, el cual puede ser visto como una optimización local en un espacio de búsqueda continuo. Se evalúan las derivadas de la función coste con respecto a los pesos de cada neurona, desde la capa de salida hasta la primera capa oculta de la red. Una vez evaluadas todas las derivadas, se realiza una corrección de los pesos mediante la resta entre el valor actual y la derivada:

$$\theta_c^{j,i} = \theta_c^{j,i} - \alpha \frac{dJ(\Theta)}{d\theta_c^{j,i}} \quad , \quad (8)$$

donde  $\alpha$  es el factor de aprendizaje o *learning rate* [32]. Este factor está relacionado con la velocidad con la cual el entrenamiento se acerca al mínimo de la función coste. Otros optimizadores actuales más eficaces que existen son *AdaBoost* o *Adam*.

Se define una época a cada vez que la red neuronal recorre todos los datos del entrenamiento. Los pesos pueden ser reajustados varias veces durante una época. El parámetro que establece cada cuántos ejemplos se ajustan los pesos se conoce como *batch size*. A mayor tamaño, se realiza un menor número de ajustes por cada época, por lo que estas serán más rápidas.

Otro concepto importante es el *dropout*, un valor que representa la probabilidad que tiene cada neurona de esa capa de desactivarse para cada *batch* (Figura 11). Se aplica el *dropout* para evitar que la red dependa mayoritariamente en un conjunto limitado de neuronas, y como consecuencia previene el sobreajuste de los datos utilizados. De esta forma se entrena la red neuronal un poco peor con los datos disponibles, para que generalice mejor a otros potenciales datos.

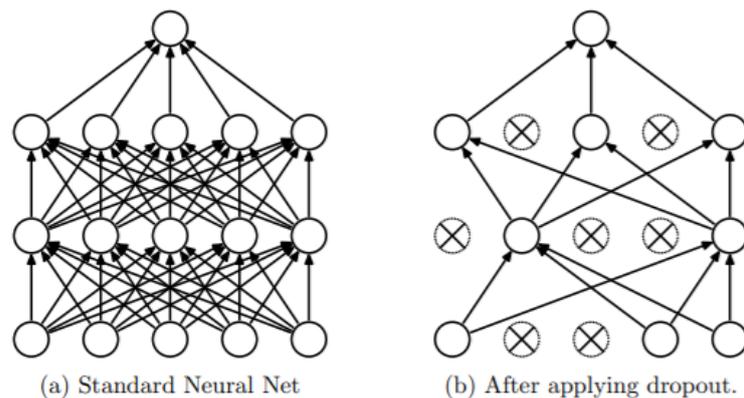


Figura 11: Ejemplo de *dropout* en una red neuronal [33].

Los elementos que rigen el proceso de entrenamiento de la red, como el tamaño del *batch size*, el número de capas ocultas o la cantidad de neuronas reciben el nombre de hiperparámetros. Para optimizar una red neuronal es necesario buscar el conjunto óptimo de hiperparámetros y de pesos. Se conocen con el nombre de hiperparámetros debido a que estos se utilizan para entrenar modelos, siendo elegidos por el científico, mientras que los parámetros se obtienen a través del proceso de entrenamiento.

## 4.2. Métricas

A la hora de evaluar los modelos de aprendizaje automático, se emplean una serie de funciones llamadas métricas [30]. Estas funciones proporcionan información sobre la capacidad de la red neuronal de discriminar entre ambos tipos de eventos. Las métricas dependen de una serie de variables que pueden ser expresadas mediante la matriz de

conteo, donde se recoge el contaje de los eventos según han sido clasificados (Cuadro 1).

	Positivos predichos	Negativos predichos
Positivos reales	TP	FN
Negativos reales	FP	TN

Cuadro 1: Matriz de confusión para un problema binario.

Los elementos de la diagonal son los verdaderos positivos y verdaderos negativos, TP y TN, estos elementos son todas las predicciones correctas de la red. Los términos no diagonales se conocen con el nombre de falsos positivos y falsos negativos, FP y FN, las predicciones erróneas de la red hacia una clase u otra. Un algoritmo será mejor cuanto mayor sean los parámetros TP y TN, o lo que es lo mismo, cuanto menor sean los valores FP y FN.

A partir de estos cuatro elementos de la matriz de confusión podemos definir dos métricas, denominadas ratios de verdaderos positivos (TPR) y falsos positivos (FPR):

$$TPR = \frac{TP}{TP + FN} \quad ; \quad (9)$$

$$FPR = \frac{FP}{FP + TN} \quad . \quad (10)$$

El ratio de verdaderos positivos corresponde al ratio de positivos que han sido acertados sobre el total de estos, mientras que el ratio de falsos positivos es la tasa de negativos reales que han sido seleccionado como positivos sobre el número total de negativos.

Los eventos de retrocesos nucleares son considerados como los eventos de la clase positiva. Debido a esto se puede definir dos nuevos parámetros, la aceptación y el poder de rechazo. Es posible establecer una relación entre el poder de rechazo y la aceptación con las métricas definidas:

$$A = TPR \quad ; \quad (11)$$

$$R_f = 1 - FPR \quad . \quad (12)$$

La función de activación en la capa de salida es una sigmoide, por lo que la salida de la red neuronal está comprendida entre 0 y 1. Debido a esto, se establece un umbral

$u$ . Todos los eventos con valores de salida que se encuentren por debajo de este umbral serán definidos como eventos de polvo (0), mientras que los eventos con valores de salida por encima serán definidos como retrocesos nucleares (1). Es necesario la búsqueda del valor óptimo de este umbral, ya que el valor  $u = 0,5$  no es el más eficaz a pesar de ser el más frecuente.

Con el fin de buscar el umbral óptimo, es necesario definir la curva ROC (Receiver Operating Characteristic). Esta curva consiste en la representación de la aceptación en función del poder de rechazo. Un buen algoritmo será aquel cuya curva ROC supere a una salida aleatoria, siendo el mejor aquel donde la aceptación y el poder de rechazo sean máximos (Figura 12).

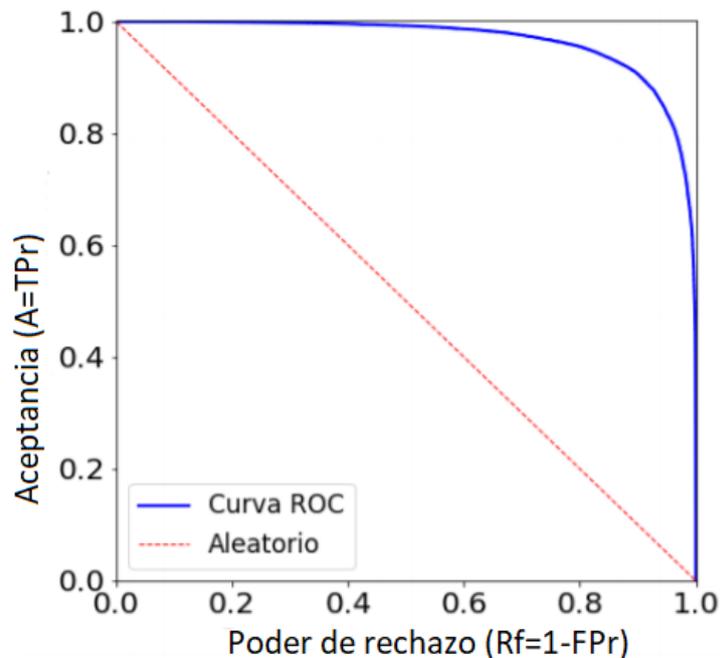


Figura 12: Ejemplo de curva ROC estándar. La curva azul representa un buen clasificador, mientras que la línea roja actúa como un mal clasificador que actúa de forma aleatoria [27].

Otra forma de evaluar un algoritmo sin necesidad de definir un umbral es mediante otra métrica conocida como el área bajo la curva ROC. En el caso perfecto, este área es la unidad, así que a mayor valor de este parámetro mejor será el algoritmo. Para el caso de una salida aleatoria (una red que no es capaz de aprender) el valor será de 0,5.

### 4.3. Fundamentos de las redes neuronales artificiales

Para el entrenamiento y evaluación de la red neuronal se han usado datos de argón, así como distintos datos de polvo de distintos tamaños, los cuales han sido simulados mediante el método de Monte Carlo generados en DEAP-3600.

Los datos se dividen en dos subconjuntos menores, donde el primero se empleará para entrenar la red neuronal mientras que el segundo se utilizará para evaluarla una vez haya terminado el proceso de entrenamiento. En este trabajo, el 75 % de los datos se han utilizado para entrenamientos, mientras que el 25 % restante se han empleado para evaluaciones. En el Cuadro 2 se muestran las composiciones de cada uno de los conjuntos de eventos.

Eventos	Eventos totales	Entrenamiento	Evaluación
Retrocesos nucleares	24973	18767	6206
Polvo 25 micras	3714	2787	927
Polvo 17 micras	3207	2402	802
Polvo 10 micras	2275	1706	569
Polvo 5 micras	429	322	107
Total de polvo	9625	7217	2405
Total	34598	25984	8611

Cuadro 2: Distribución de los 34598 eventos entre conjuntos de entrenamiento y evaluación según su procedencia.

Por cada evento se tienen 255 valores de entrada, uno por cada PMT del detector, que son la carga recolectada por cada uno de los fotomultiplicadores, normalizada a la carga total del evento:

$$\bar{q}_i = \frac{q_i}{\sum_i q_i} \quad , \quad (13)$$

donde  $q_i$  es la luz en fotoelectrones recolectada por cada  $i$ -ésimo PMT,  $\bar{q}_i$  el valor normalizado. Esta normalización es necesaria para poder imponer que la red neuronal aprenda patrones de luz, y no de la energía total del evento.

La capa de salida está compuesta por una neurona, cuya función de activación es la sigmoide para poder realizar la clasificación binaria, mientras que para las capas ocultas la función de activación es la función ReLu ( $h(z) = \max(0, z)$ ). En la Figura 13 se observan ejemplos gráficos de estas dos funciones de activación.

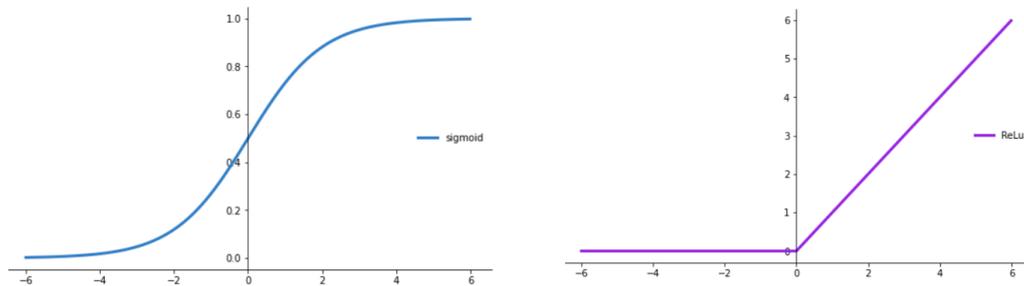


Figura 13: Representación de dos funciones de activación. A la izquierda se representa la función sigmoïdal y a la derecha la función ReLu.

Un estudio previo de diferentes funciones de activación mostró que la función ReLu proporciona los mejores resultados en menor tiempo de entrenamiento. También se probaron varios optimizadores, siendo el más rápido el optimizador Adam. Este optimizador adapta el ratio de aprendizaje en función de cómo estén distribuidos los parámetros. Si los parámetros están muy dispersos, el ratio de aprendizaje aumentará [30].

El trabajo se ha realizado mediante dos librerías de Python, una es especializada en redes neuronales, como *Keras* [34], y una segunda especializada en aprendizaje automático, como *scikit-learn* [35].

#### 4.4. Fundamentos de las Redes Neuronales Convolucionales

Una Red Neuronal Convolutiva (CNN) es un tipo de red neuronal especializada para procesar datos, con una topología similar a una cuadrícula [30]. Los ejemplos que se incluyen son las series de datos temporales, donde puede ser pensado como una cuadrícula 1D que está tomando datos en intervalos regulares de tiempo. También pueden procesar imágenes, pensadas como cuadrículas de píxeles en 2D.

En las CNN la red toma como entrada los píxeles de una imagen. Se emplea una operación matemática llamada convolución, cuya expresión es:

$$s(t) = (x.w)(t) \quad , \quad (14)$$

donde  $x$  es la entrada y  $w$  una matriz de tamaño menor que la imagen de entrada, llamada *kernel* o filtro. La salida,  $s$ , se conoce como mapa de características. La variable  $t$  es el tiempo.

En el proceso de convolución, se realiza el producto escalar entre el *kernel* y un conjunto

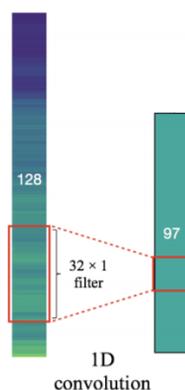


Figura 14: Ejemplo de *kernel* convolutivo en una dimensión.

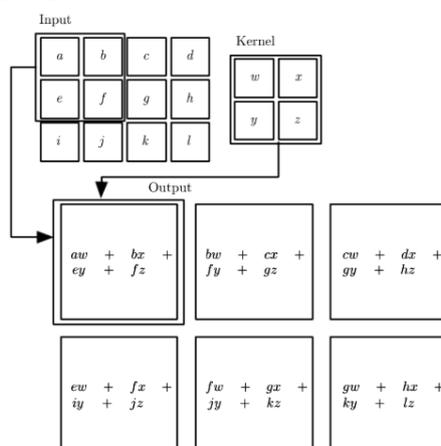


Figura 15: Ejemplo de *kernel* convolutivo en dos dimensiones

de píxeles cercanos de la entrada. A su vez, el *kernel* se desplaza a lo largo de la imagen hasta que la cubre completamente (Figuras 14 y 15). Se genera así una nueva matriz de salida, siendo esta una nueva capa de neuronas ocultas. Se conoce con el nombre de *zancada* al tamaño de desplazamiento del *kernel*.

En una red neuronal tradicional, cada unidad de entrada interactúa con cada unidad de salida. Por otra parte, en una red convolucional las interacciones son dispersas, almacenándose menos parámetros. En general, en una CNN se necesitan menos operaciones para obtener la salida de la red, mejorando la memoria y la eficiencia. Durante el proceso convolutivo, se aplican varios *kernels*.

Después de cada capa de la CNN se aplica un *pooling* para prevenir el overfitting, es decir, para evitar un sobreprocesamiento. El parámetro *Max - pooling* reduce

el tamaño de las imágenes filtradas prevaleciendo las características más importantes detectadas por cada filtro (Figura 16). Esta reducción será proporcional al valor del *Max – pooling*.

7	4	2	5		
1	2	3	4	7	5
5	3	3	2	5	3
2	4	1	2		

Figura 16: Ejemplo de *Max – Pooling* 4 para un *kernel* de 2x2 y zancada 2 [30]. Se reduce la matriz un factor 4, el valor del *Max – Pooling*.

En una red neuronal convolucional la capa de salida estará formada por dos neuronas, una por cada clase que se clasifica. Además, la función de activación de esta capa de salida será la función sigmoide. Para las capas ocultas se empleará la función de activación ReLu. El conjunto formado el *kernel*, el *pooling* y la función de activación recibe el nombre de bloque convolutivo.

## 5. Análisis con Redes Neuronales

En esta sección se estudian todos los resultados obtenidos mediante el algoritmo de redes neuronales de perceptrón multicapa (ANN).

### 5.1. Datos desbalanceados

Con el fin de realizar una buena clasificación entre los eventos de polvo y los eventos de retrocesos nucleares, es necesario buscar patrones que caractericen a ambos tipos de eventos. Para ello se realiza un histograma donde se incluyen todos los valores de entrada de todos los eventos de polvo y de retrocesos nucleares de argón. En la Figura 17 se representa el número de cuentas medidas para todos los valores de carga normalizada medidos por los PMTs del detector.

Para las muestras de polvo, en promedio, la carga total del evento se focaliza sobre un número menor de PMTs, debido a la sombra que generan los granos de polvo. En los eventos de retrocesos nucleares la luz está más distribuida, recogiendo cada PMT un porcentaje menor de carga normalizada. Estas diferencias distintivas en los patrones de ambos tipos de eventos serán explotadas por la red neuronal para clasificarlos.

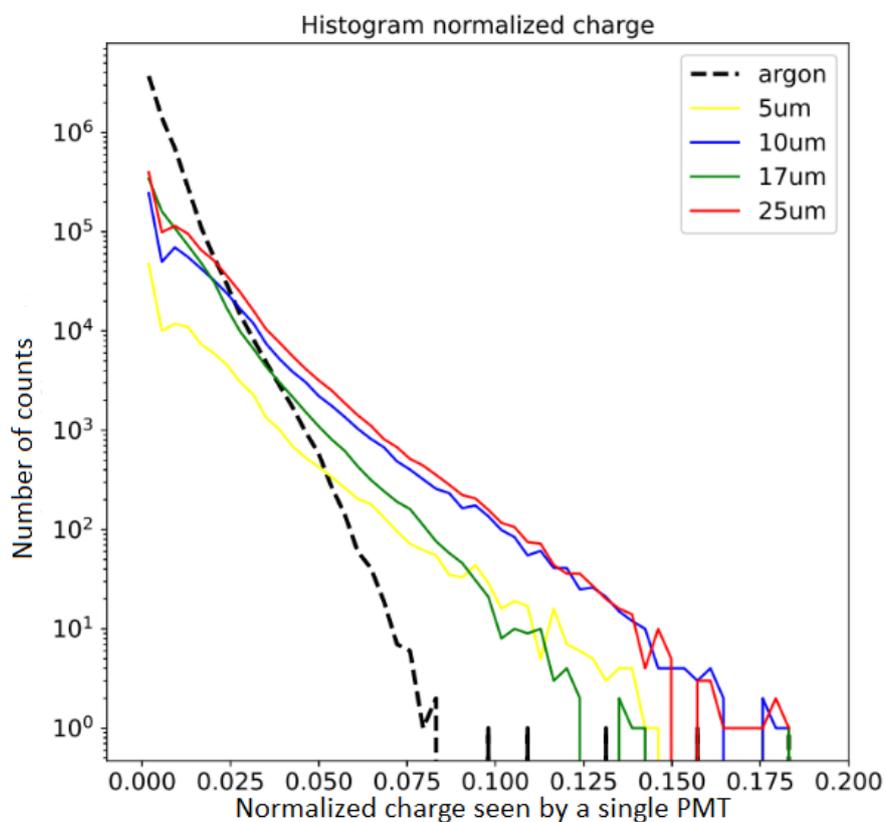


Figura 17: Número de cuentas medidas para cada carga normalizada recibida por cada PMT del detector. Se toman todos los datos de cada muestra de polvo y de los retrocesos nucleares de argón.

Una vez analizadas las diferencias entre ambos tipos de eventos, se intenta entrenar una red neuronal tipo ANN usando exclusivamente las muestras de polvo de 25 micras. Esta red neuronal posee una estructura de 2 capas ocultas, con 256 neuronas en la primera capa y 64 en la segunda. El *dropout* establecido es de 0,10, junto a un *batch size* de 32 y 8 épocas. En la Figura 18 se observan las distribuciones a la salida de la red neuronal para los eventos de evaluación de retrocesos nucleares (azul) y los eventos de polvo (naranja).

Los eventos de argón se etiquetan como 1 y los eventos de polvo se etiquetan como 0. En una red neuronal bien entrenada debería reproducirse, es decir, se tendrían que catalogar correctamente los eventos. A partir de los resultados que se obtienen a la salida de la red neuronal, se puede comprobar que la red no es capaz de distinguir entre ambos tipos de eventos. Si la red fuese capaz de diferenciarlos, los eventos de polvo serían etiquetados

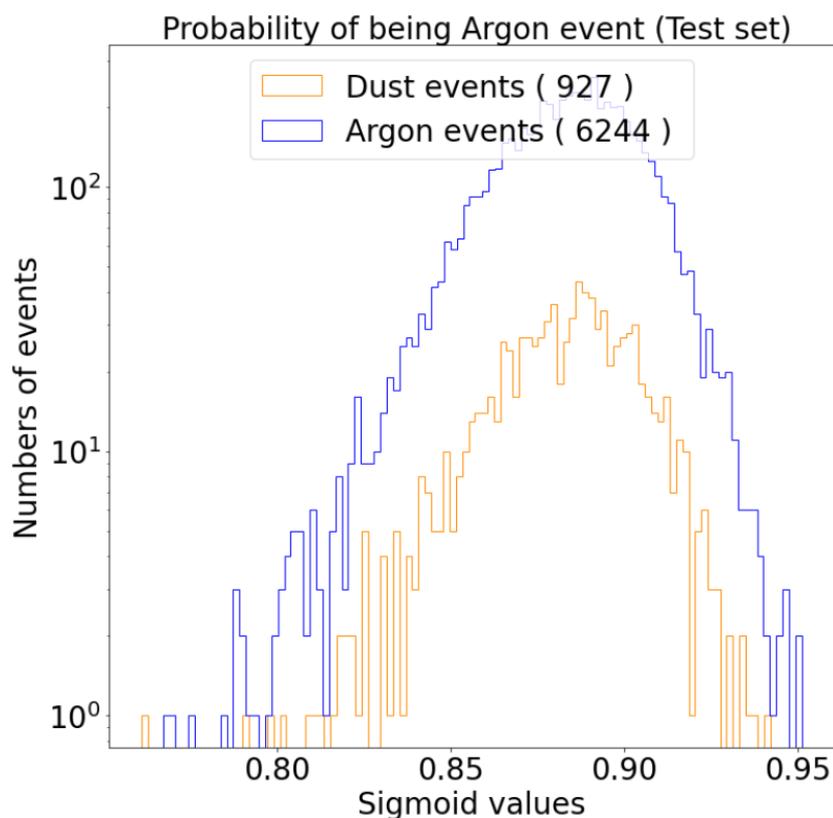


Figura 18: Distribución del resultado de la red para las poblaciones de retrocesos nucleares y eventos de polvo que forman el conjunto de test.

con valores próximos a 0 en la sigmoide y los eventos de argón en las proximidades de 1. Por el contrario, observamos que ambos tipos de eventos se etiquetan mayoritariamente en el rango  $[0,75, 0,95]$ , siendo indistinguibles para la red neuronal. Por lo tanto, este algoritmo es ineficiente para separar fondo y señal.

Otro ejemplo que muestra el bajo rendimiento de la red es su curva ROC (Figura 19), que evidencia que la clasificación es análoga a una clasificación aleatoria. Para poder establecer un poder de rechazo de los eventos de fondo del 98,9 % tendríamos que establecer el umbral  $u = 0,925$ , teniendo una aceptación de retrocesos nucleares del 0,8 %.

Para mejorar el rendimiento del algoritmo, se propone balancear los datos. Cuando existen datos desbalanceados, la red neuronal tiende a favorecer a la clase con mayor proporción de observaciones, fenómeno que se observa en la Figura 18. Esto debido a la cantidad de datos insuficientes de muestras de polvo. Resolver el problema del desbalance

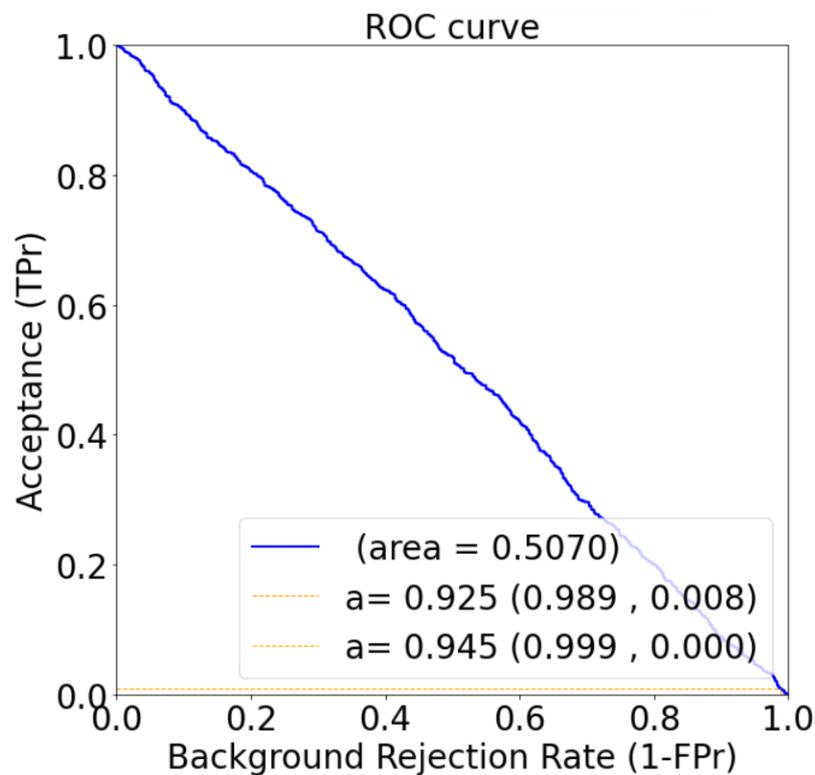


Figura 19: Curva ROC correspondiente a la red analizada.

de datos permitiría solucionar el mal entrenamiento de la red y la nula capacidad de distinguir ambos tipos de eventos.

## 5.2. Datos balanceados con SMOTE

Para tratar el problema de desbalance entre ambas clases de eventos, emplearemos una técnica del sobremuestreo (over-sampling), conocida como SMOTE (Synthetic Minority Over-Sampling Technique) [36]. El funcionamiento de esta técnica consiste en generar nuevas muestras interpoladas de la clase minoritaria.

Los nuevos valores sintéticos son generados a partir de los  $k$  eventos vecinos más cercanos de un evento de la clase minoritaria. SMOTE genera de forma aleatoria un número entre 1 y  $k$  para escoger qué vecino se toma como referencia. Una vez tomado este número, se genera un segmento entre la muestra minoritaria actualmente procesada y el vecino correspondiente al número aleatorio. En este segmento se crea el nuevo valor sintético, cuya posición en él será aleatoria. Todo este procedimiento se repite hasta obtener el mismo número de datos de ambos tipos (Figura 20). En este trabajo se ha

usado una configuración basada en los 5 vecinos más próximos.

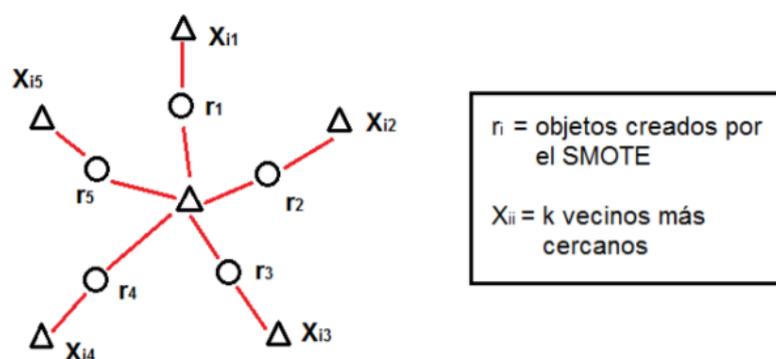


Figura 20: Ejemplo visual del funcionamiento de SMOTE. Los valores sintéticos  $r_i$  son generados mediante interpolación entre la muestra actualmente procesada (término central) y los  $k$  vecinos más cercanos  $X_{ii}$  [37].

Para obtener los datos sintéticos, el empleo de SMOTE se debe de realizar después de aplicar la división de datos de entrenamiento y datos de test (set). Realizarlo al revés causaría que la red se autoengañe, ya que es posible que dos eventos muy parecidos acaben por separado en el train y en el test. Esto provoca que a la hora de evaluar, la red analice eventos que ya se han analizado previamente en el entrenamiento. Hay que tener en cuenta que debido al coste computacional necesario para generar los datos, no es posible tener una cantidad suficiente de ellos. Sólo se dispone de una cantidad para poder realizar una evaluación rápida.

Una vez aplicado SMOTE sobre el *dataset* original en el código de la red neuronal, se representan los resultados obtenidos para las muestras de polvo de 25 micras (Figuras 21 y 22). Se observan así las diferencias con respecto al caso anterior con datos desbalanceados.

Estos nuevos resultados muestran que la red estaría funcionando mucho mejor que en caso anterior sin SMOTE, siendo capaz de diferenciar ambos tipos de eventos. En la salida de la red neuronal, siendo *score* el pronóstico de la red, se observa cómo ambas distribuciones presentan máximos en sus etiquetas correspondientes. También se obtiene una gran mejora de la curva ROC, la esperada para un buen clasificador. Para un poder de rechazo de eventos de fondo del 90 % y del 50 % tendremos aceptancias de retrocesos nucleares del 36,7 % y del 88,5 % respectivamente.

Los resultados obtenidos son prometedores. Gracias a la técnica de SMOTE, produ-

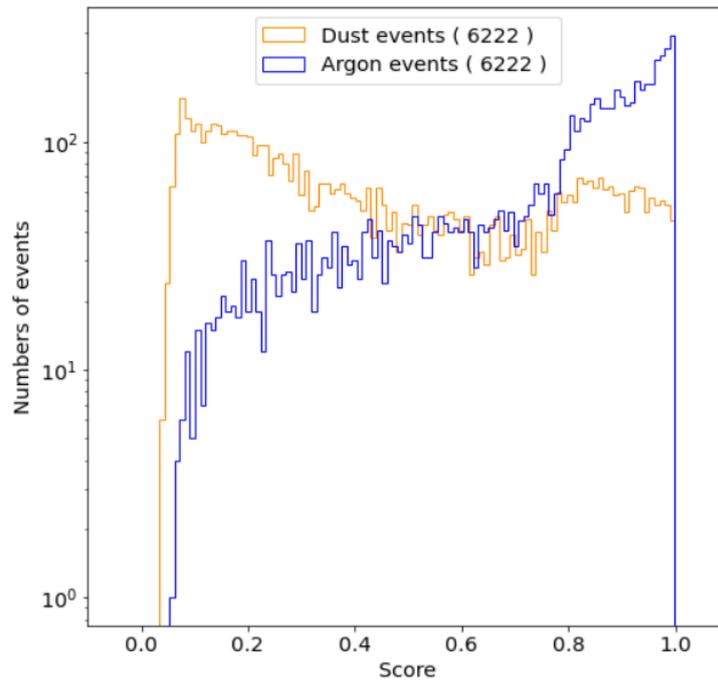


Figura 21: Distribución del resultado de la red para las poblaciones de retrocesos nucleares y eventos de polvo de 25 micras que forman el conjunto de test tras aplicar SMOTE.

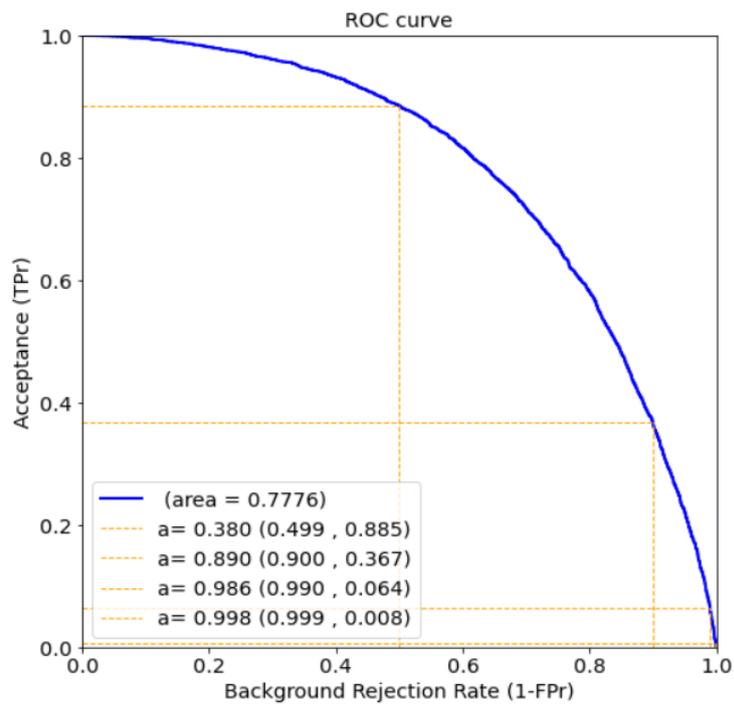


Figura 22: Curva ROC de la ANN tras aplicar SMOTE al conjunto de test de eventos de polvo de 25 micras.

ciendo un balance de datos de la muestra minoritaria mediante interpolación, el mal funcionamiento de la red ha sido subsanado. Aparte de la implementación de SMOTE también se produjo un cambio en los hiperparámetros una vez visto que el algoritmo era capaz de clasificar eventos. Los resultados que se muestran son con los valores óptimos de los hiperparámetros. El proceso de búsqueda de esos valores óptimos se muestra en el siguiente apartado.

### 5.3. Optimización hiperparámetros

Se define el parámetro ECM como el error cuadrático medio entre las predicciones y los valores de las muestras de los tests. La expresión de este parámetro es:

$$ECM = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad , \quad (15)$$

donde  $n$  es el número total de eventos de test,  $\hat{Y}_i$  es el valor  $i$ -ésimo de las predicciones y  $Y_i$  es el  $i$ -ésimo valor del test. Un algoritmo será mejor cuanto menor sea el error cuadrático medio.

Para maximizar el rendimiento de la red, es necesario realizar una optimización de hiperparámetros, obteniendo así el menor error cuadrático medio posible. Para ello, el procedimiento seguido consiste en realizar una búsqueda codiciosa. Primero, se fijan todos los hiperparámetros excepto uno, que se modifica entre varios valores hasta lograr el menor ECM posible. Este ya se fija definitivamente y se procede con el resto.

Los hiperparámetros a optimizar son el número de neuronas por capa, el *batch size* y el *dropout*. El número de épocas se mantendrá fijo con 7, realizándose 15 repeticiones en cada distribución de hiperparámetros. Se obtiene en cada caso el error cuadrático medio, realizando al final la media y la desviación estándar. Las distribuciones con los valores más bajos serán las elegidas. El primer parámetro a optimizar es el número de neuronas por capa (Cuadro 3). Se mantiene fijo el *dropout* a 0,15 y el *batch size* a 32.

La arquitectura con mejores resultados es aquella con 256 y 32 neuronas en la primera y segunda capa oculta respectivamente, ya que proporciona los valores más bajos. También se obtiene que la mejor red debe estar compuesta por dos capas ocultas. Al añadir una tercera capa oculta, los resultados empeoran. Una vez fijada esta estructura, se optimiza el *batch size*, como se observa en el Cuadro 4.

Primera capa	Segunda capa	Tercera capa	Media ECM	Desv. estand ECM
128	16	0	0,2225	0,0095
128	32	0	0,2200	0,0081
128	64	0	0,2232	0,0069
256	16	0	0,2179	0,0094
256	32	0	0,2154	0,0044
256	64	0	0,2237	0,0131
256	32	16	0,2284	0,0107
512	16	0	0,2214	0,0151
512	32	0	0,2162	0,0060
512	64	0	0,2275	0,0173

Cuadro 3: Distintas arquitecturas junto a su media y desviación estándar de los distintos ECM para un *dropout* de 0,15, un *batch size* de 32 y 7 épocas con 15 repeticiones.

Batch size	16	32	64
Media ECM	0,2216	0,2168	0,2146
Desv. estand ECM	0,0076	0,0067	0,0091

Cuadro 4: Distintos *batch size* junto a su media y desviación estándar de los distintos ECM para un *dropout* de 0,15, una arquitectura de 256 y 32 neuronas en dos capas ocultas y 7 épocas con 15 repeticiones.

El *batch size* estudiado con los mejores resultados que se obtiene es el de 64. Se fija este parámetro y se procede a la búsqueda del *dropout* que minimice el ECM. Como se observa en el Cuadro 5, este valor es de 0,20.

Así pues, gracias a este estudio minucioso, la mejor red encontrada es aquella que posea los hiperparámetros mostrados en el Cuadro 6. Esta configuración proporciona el menor error cuadrático medio.

#### 5.4. Rendimiento para cada muestra de polvo

En este apartado se procede a mostrar los resultados obtenidos tras aplicar SMOTE y optimizar los distintos hiperparámetros en cada una de las distintas muestras de polvo de forma independiente, observando las diferencias entre cada una de ellas. Los resultados de las muestras de polvo de 25 micras fueron descritas previamente para mostrar el buen funcionamiento de SMOTE.

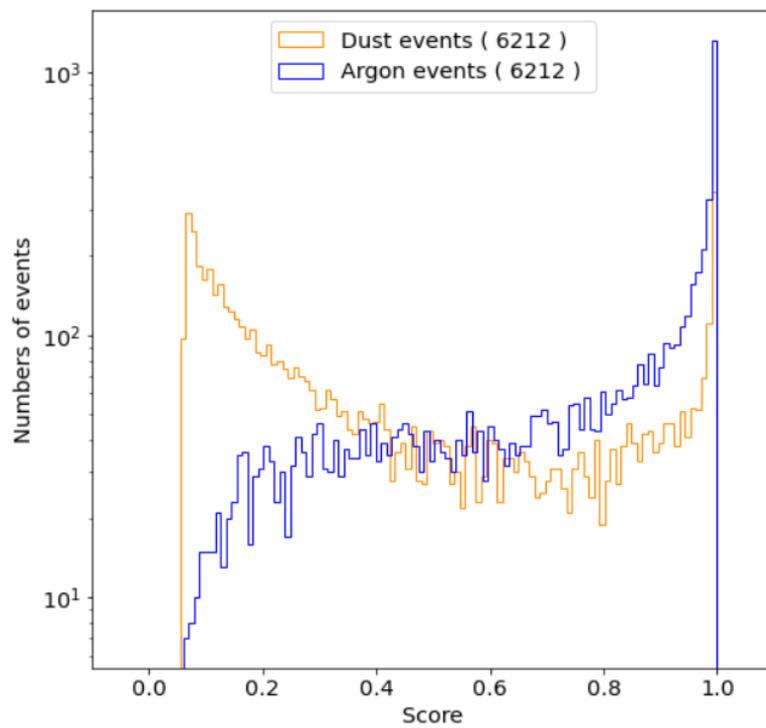


Figura 23: Distribución del resultado de la red para las poblaciones de retrocesos nucleares y eventos de polvo de 17 micras que forman el conjunto de test tras aplicar SMOTE.

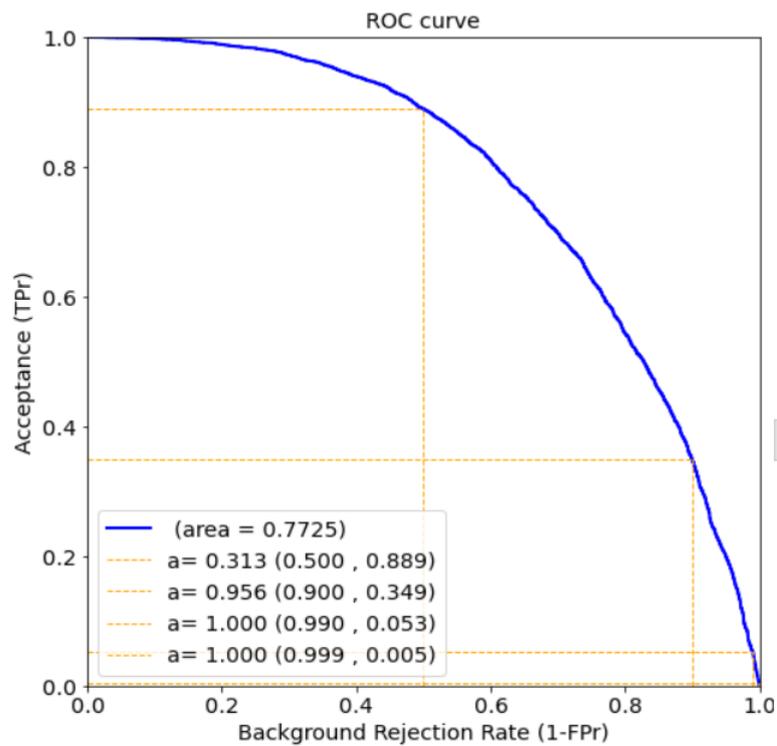


Figura 24: Curva ROC de la ANN tras aplicar SMOTE al conjunto de test de eventos de polvo de 17 micras.

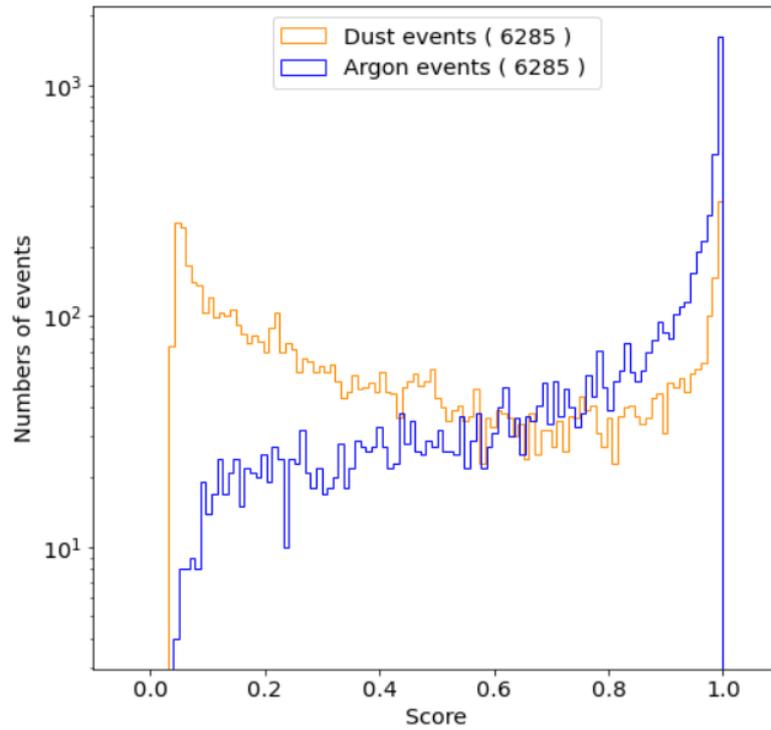


Figura 25: Distribución del resultado de la red para las poblaciones de retrocesos nucleares y eventos de polvo de 10 micras que forman el conjunto de test tras aplicar SMOTE.

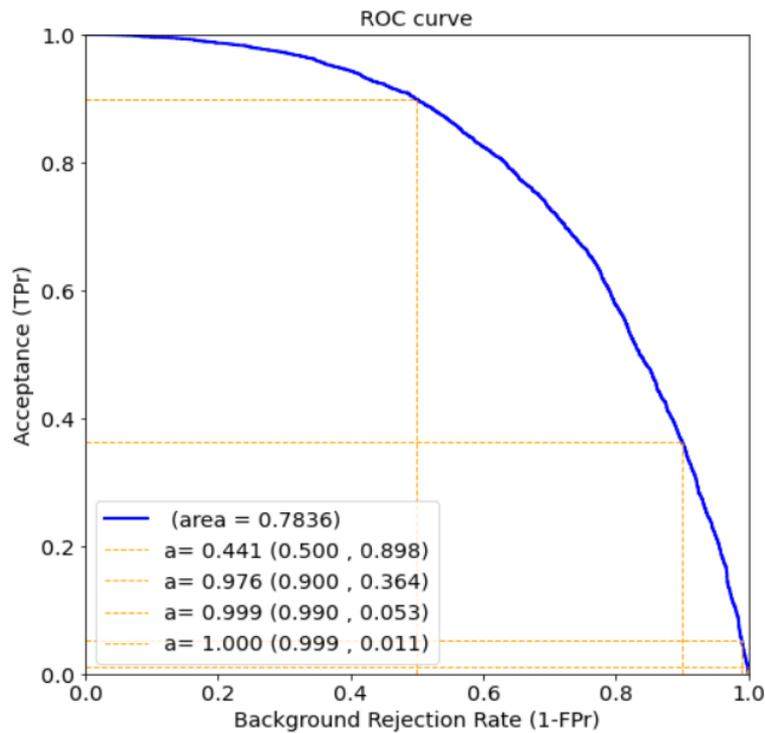


Figura 26: Curva ROC de la ANN tras aplicar SMOTE al conjunto de test de eventos de polvo de 10 micras.

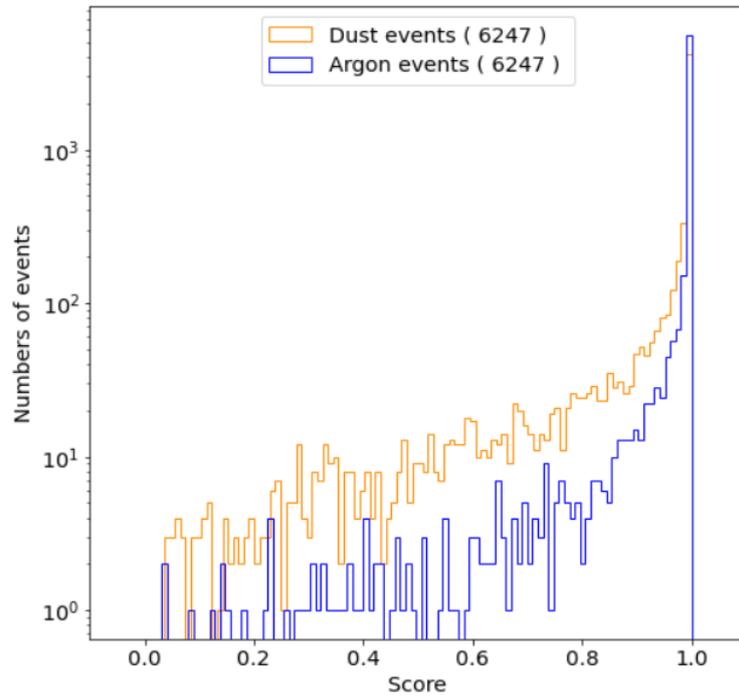


Figura 27: Distribución del resultado de la red para las poblaciones de retrocesos nucleares y eventos de polvo de 5 micras que forman el conjunto de test tras aplicar SMOTE.

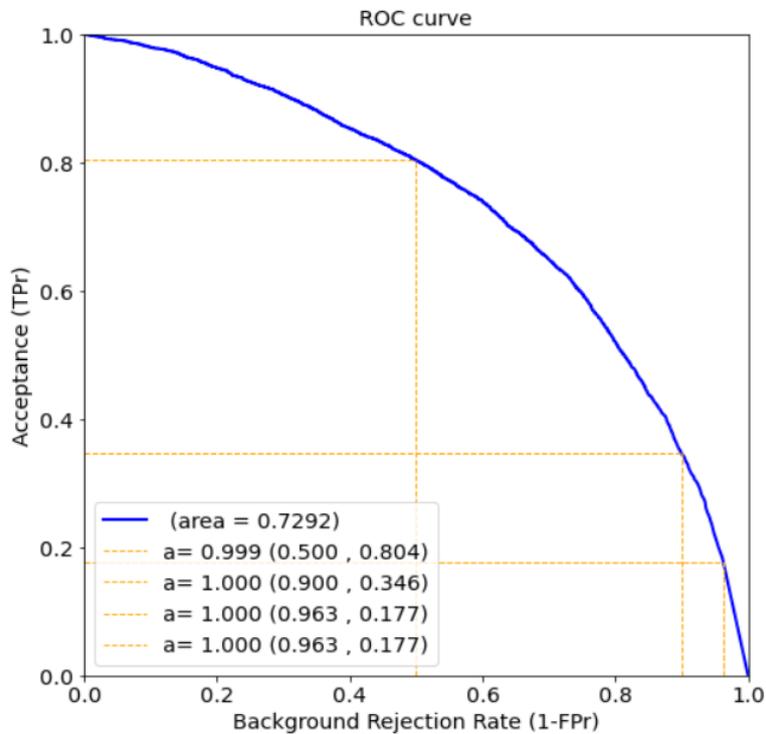


Figura 28: Curva ROC de la ANN tras aplicar SMOTE al conjunto de test de eventos de polvo de 5 micras.

Dropout	0,10	0,15	0,20
Media ECM	0,2247	0,2168	0,2073
Desv. stand ECM	0,0135	0,0047	0,0098

Cuadro 5: Distintos *dropouts* junto a su media y desviación estándar de los distintos ECM para un batch size de 64, una arquitectura de 256 y 32 neuronas en dos capas ocultas y 7 épocas con 15 repeticiones.

Hiperparámetro	Valor
Neuronas primera capa	256
Neuronas segunda capa	32
Batch Size	64
Dropout	0,20

Cuadro 6: Hiperparámetros estudiados que proporcionan el menor error cuadrático medio a la red neuronal.

A través de las gráficas mostradas se puede determinar un buen funcionamiento de la red para las muestras de 25 (Figuras 17 y 18), 17 (Figuras 23 y 24) y 10 (Figuras 25 y 26) micras, mientras que para 5 micras (Figuras 27 y 28) se determina que la red no es capaz de diferenciar ambos tipos de eventos. Esto es debido a la existencia de muy pocos datos de 5 micras, SMOTE crea datos que, aunque sean muy parecidos, producen las suficientes diferencias en el comportamiento para que sean catalogados como retrocesos nucleares. A pesar que la red es capaz de clasificar eventos, esta clasificación no es lo suficientemente buena. Es necesario un aumento de los datos de las muestras de 5 micras para sacar mejores conclusiones.

Con respecto a las muestras de polvo de 25, 17 y 10 micras, sólo se perciben pequeñas diferencias en el comportamiento. Existe una pequeña mejora en los resultados de 17 y 10 micras, producidas posiblemente por procesos estocásticos, provocando una pequeña mejora en la desviación estándar. Todas las distribuciones presentan máximos en sus etiquetas correspondientes. También se obtienen buenas curvas ROC, muy similares entre todas las muestras de polvo. Estas curvas ROC son las esperadas para un clasificador de eventos aceptable.

### 5.5. Rendimiento para todas las muestras de polvo juntas

En este apartado se muestran los resultados obtenidos tras aplicar SMOTE y optimizar los distintos hiperparámetros para las muestras de polvo de forma conjunta. En la realidad existirá una mezcla de todos los tamaños, por lo que es importante realizar un análisis sobre ello.

Para el caso de las cuatro muestras de polvo de distintos tamaños, a pesar del buen funcionamiento de la red, los resultados son peores. Las muestras de polvo de distintos tamaños poseen distintas distribuciones. Al interpolar entre todas ellas, algunos eventos sintéticos generados por SMOTE pueden tomar valores que corresponden a los retrocesos nucleares. Esto provoca que el algoritmo confunda eventos, generando así un empeoramiento de la salida de la red (Figura 29) y en su curva ROC (Figura 30). Para unos poderes de rechazo del 90 % y del 50 % los valores de la aceptación son del 20,6 % y del 78,8 % respectivamente.

Para los mismos poderes de rechazo, los valores de aceptación de retrocesos nucleares son menores en comparación con las muestras de polvo de forma independiente. Si la aceptación de datos es demasiada baja, es posible que se rechacen eventos potencialmente interesantes.

### 5.6. Clasificación en función de posición reconstruida

En este apartado se buscan patrones en los eventos que han sido catalogados como “TP”, “FP”, “TN” y “FN”. Se analiza el aprendizaje que está realizando la red y qué patrones clasifica.

Para los poderes de rechazo del 99 % y del 99.9 %, las aceptancias de retrocesos nucleares son demasiado bajas. No se posee una cantidad de datos significativa. Por este motivo, sólo se tendrán en cuenta los poderes de rechazo del 50 % y del 90 %. Para estos dos casos, se representan las distribuciones en la posición Z reconstruida para todas las 4 muestras de polvo juntas, como se observa en las Figuras 31 y 32. Sólo se emplea el conjunto test de todas las muestras de polvo debido a que en la realidad se tendrá una mezcla de todos los tamaños de polvo. Cabe destacar que una vez entrenada la red, los eventos sintéticos generados por SMOTE no se tienen en cuenta en este apartado. Sólo se emplean los datos de test originales. Este factor implica que para un poder de rechazo del 50 %, no se tenga la mitad de eventos de polvo correctamente etiquetados. Es posible que la clasificación de eventos sea más efectiva para los eventos sintéticos o para los originales.

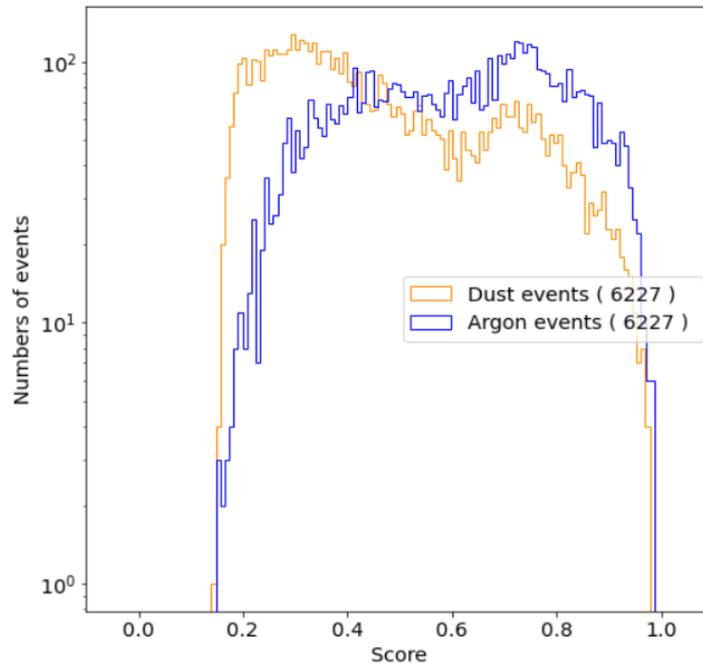


Figura 29: Distribución del resultado de la red ANN para las poblaciones de retrocesos nucleares y eventos de polvo de 25, 17, 10 y 5 micras que forman el conjunto de test tras aplicar SMOTE.

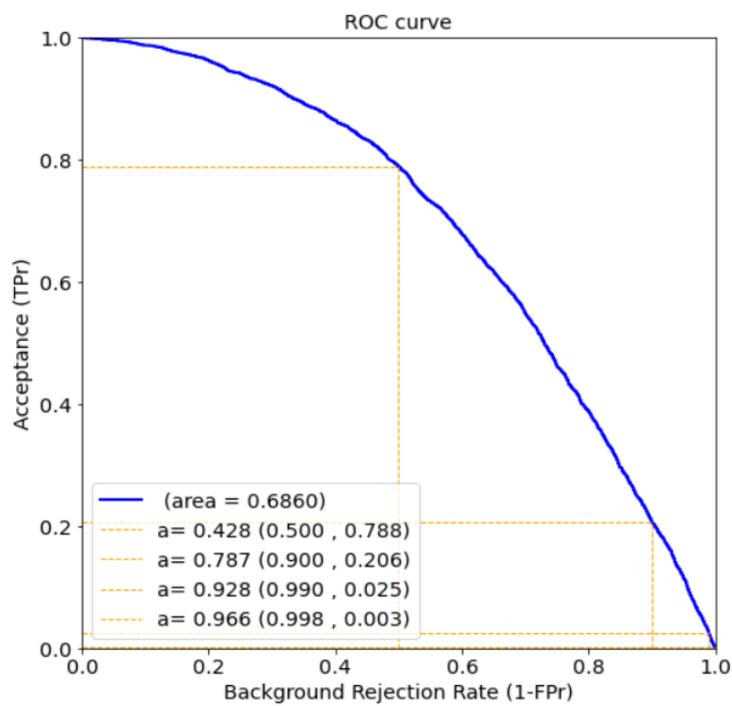


Figura 30: Curva ROC de la ANN tras aplicar SMOTE al conjunto de test de eventos de polvo de 5, 10, 17 y 25 micras.

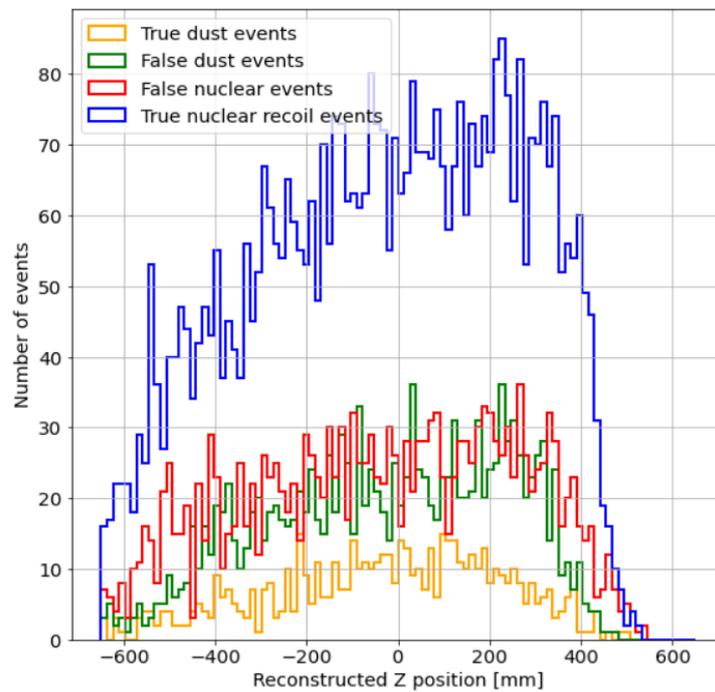


Figura 31: Histograma de posición reconstruida para un poder de rechazo del 50% sobre el conjunto datos de test original de todas las muestras de polvo.

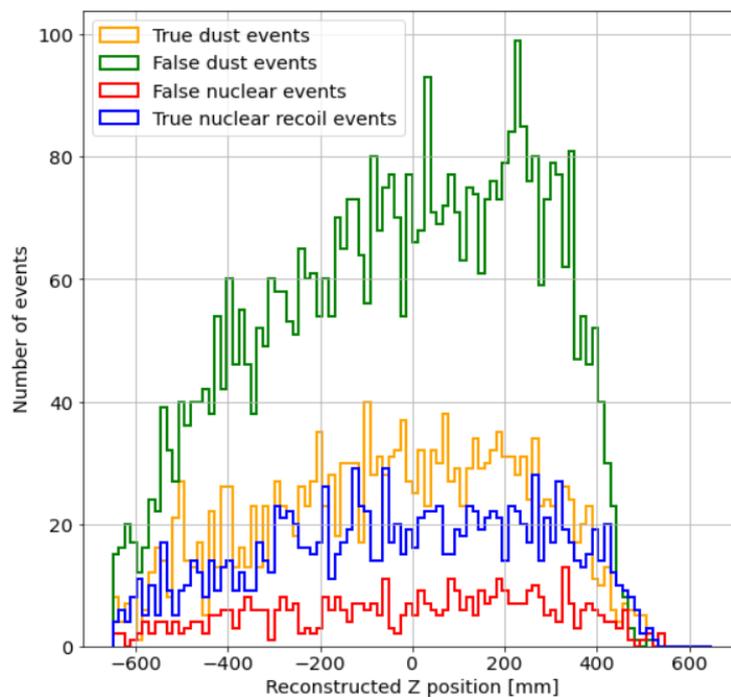


Figura 32: Histograma de posición reconstruida para un poder de rechazo del 90% sobre el conjunto datos de test original de todas las muestras de polvo.

Para un poder del rechazo del 50 % (Figura 31), predominan de retrocesos nucleares correctamente clasificados (true nuclear recoil events) frente a los retrocesos erróneamente clasificados, es decir, eventos de retrocesos catalogados como polvo (false dust events). De forma contraria, existe una mayor cantidad de eventos de polvo clasificados erróneamente, eventos de polvo catalogados como retrocesos nucleares (false nuclear events), que correctamente (true dust events).

Este resultado es explicable mediante el umbral establecido en la Figura 30,  $u = 0,428$ . Todos los eventos a la derecha del umbral son catalogados como retrocesos nucleares, mientras que todos los que están a la izquierda se catalogan como eventos de polvo. Debido a la distribución de los eventos en la Figura 29, se justifica el gran acierto de eventos de retrocesos nucleares y el predominio de errores de eventos de polvo.

Cuando el poder de rechazo es del 90 % (Figura 32), sucede una situación diferente. Predominan los errores de retrocesos nucleares y los aciertos de eventos de polvo. En este caso, el umbral establecido es  $u = 0,787$ , está situado más a la derecha que el caso anterior. Existirán más eventos de cola de retrocesos nucleares que serán erróneamente clasificados y se realiza una mejor clasificación de los eventos de polvo. Es un peor clasificador de eventos.

Existe una mayor cantidad de eventos en el hemisferio superior del detector, aunque no es destacable en comparación con el hemisferio inferior. En las partes superiores e inferiores del detector, no se detectan apenas eventos. El algoritmo no es capaz de distinguir ninguna propiedad en ambos tipos de eventos. No es posible extraer ningún patrón característico. Ambos tipos de eventos poseen las mismas distribuciones de posición reconstruída.

La red neuronal es capaz de clasificar mejor los eventos sintéticos de polvo. Cuando el poder de rechazo es del 50 %, la red no es capaz de clasificar los eventos originales, mientras que para el 90 % sí. Se clasifican correctamente el 24,1 % y el 79,9 % de los eventos originales para ambos poderes de rechazo respectivamente. Los valores de aceptación de retrocesos nucleares son los obtenidos por las curvas ROC.

### 5.7. Clasificación en función de energía

Una vez analizado el patrón de la posición reconstruída, se procede a analizar el patrón de energía. De la misma forma, se estudian los poderes de rechazo del 50 % y del 90 %. Se representan las distribuciones de la energía para todas las 4 muestras de polvo juntas (Figuras 33 y 34). Igualmente, sólo se analiza el conjunto de test original una vez la red

ha sido entrenada.

Para un poder de rechazo del 50 % (Figura 33) predominan los aciertos de retrocesos nucleares y los errores de polvo. Para un poder de rechazo del 90 % (Figura 34) predominan los aciertos de polvo y los errores de retrocesos nucleares. La justificación es la vista anteriormente, el umbral.

Se puede determinar un patrón característico en ambas distribuciones. La red es capaz de separar mucho mejor los eventos de retrocesos nucleares de los de polvo cuando se producen a menor energía. La red comete errores y clasifica los eventos con mucha luz como polvo, mientras que los eventos con poca luz los clasifica como retrocesos nucleares. Lo esperable es no obtener ningún patrón. La red sobreexplota esta característica y no clasifica los eventos de forma óptima.

### 5.8. Validación sobre datos de $^{39}\text{Ar}$

La red neuronal es entrenada sobre simulaciones porque no se posee la suficiente estadística, y porque no es posible asegurar que cada dato experimental esté correctamente etiquetado. Es importante asegurarse que el algoritmo es aplicable a datos experimentales.

Para verificar el modelo de aprendizaje automático, se necesitan datos reales medidos a partir de retrocesos electrónicos. Una manera de hacerlo es buscar un conjunto de datos experimentales que sea abundante y se tenga bien clasificado ( $^{39}\text{Ar}$ ). Se analiza sólo si el comportamiento es el mismo o no. La red no podría ser implementada si no es capaz de distinguir entre eventos reales y simulados. El conjunto de datos de argón-39 está formado por 100874 eventos de datos reales del detector y 17484 eventos de datos simulados.

En la Figura 35 se observan las distribuciones normalizadas, tanto para los datos simulados como los datos reales, aplicando la misma red neuronal una vez entrenada con las simulaciones. Se muestra que ambas distribuciones son muy similares entre ellas, lo cual implica que la red no distingue entre realidad y simulación. Es posible implementar la red para analizar datos reales.

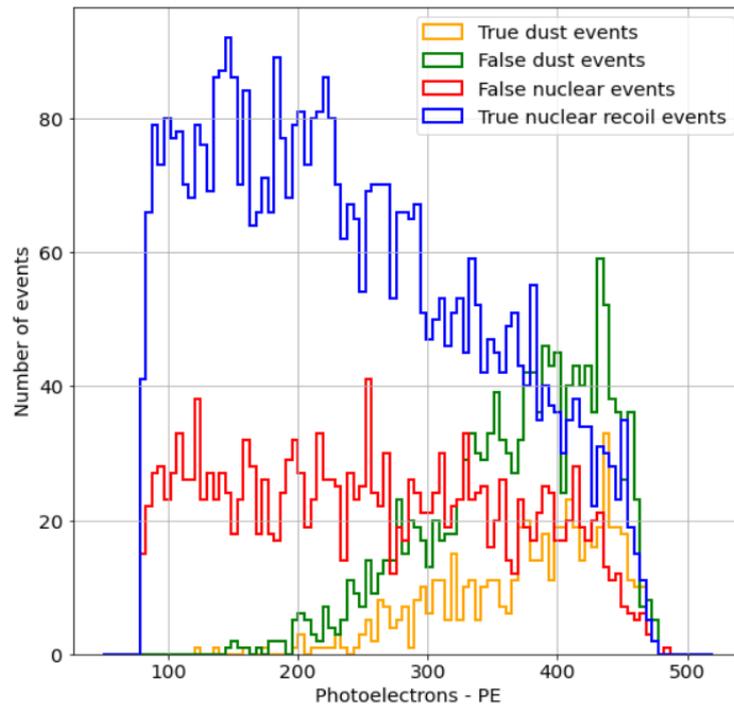


Figura 33: Histograma de energía para un poder de rechazo del 50% sobre el conjunto datos de test original de todas las muestras de polvo.

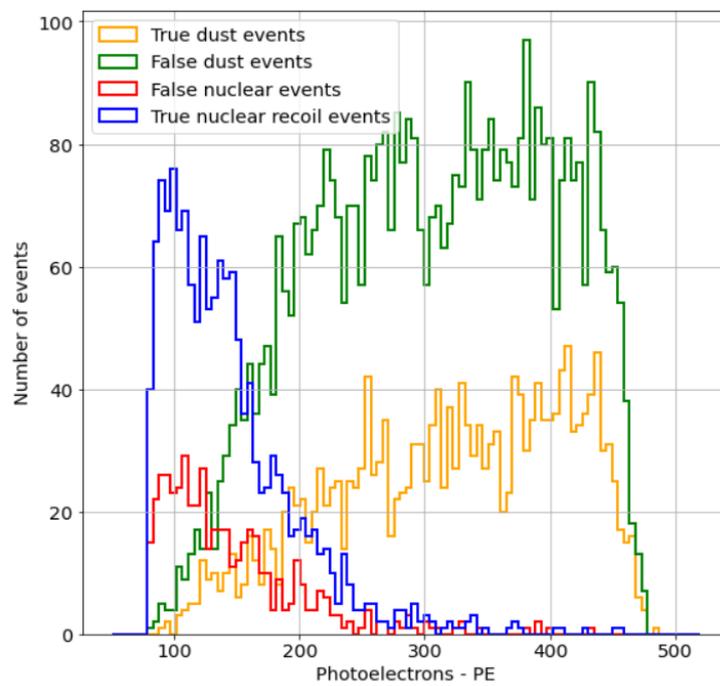


Figura 34: Histograma de energía para un poder de rechazo del 90% sobre el conjunto datos de test original de todas las muestras de polvo.

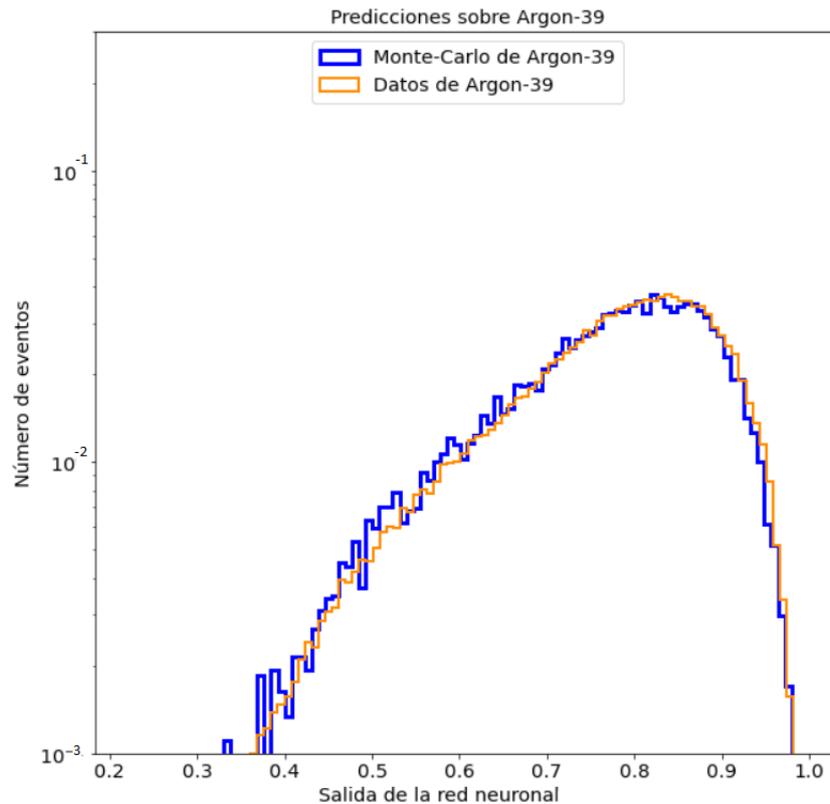


Figura 35: Aplicación de la mejor red tipo ANN sobre datos de retrocesos electrónicos experimentales y simulados, producidos por el decaimiento beta del  $^{39}\text{Ar}$ . Ambos subconjuntos están normalizados para poner énfasis en las distribuciones.

## 6. Análisis con Redes Neuronales Convolucionales

### 6.1. Optimización de los hiperparámetros

Mediante un análisis exhaustivo, se determinó que los hiperparámetros no afectan al desarrollo de la red neuronal convolucional. Los únicos términos relevantes que influyen en el comportamiento son el tamaño del *kernel*, el número de filtros, y el *pooling*. El número de capas de convolución también es estudiado.

Por falta de espacio sólo se introducen los valores optimizados de los parámetros en el Cuadro 7.

### 6.2. Rendimiento de la CNN

Se introducen los resultados para todas las muestras de polvo juntas, debido a que en la realidad se tiene una distribución de los diferentes tamaños de polvo. En primer lugar

Variable	Valor
Número de capas de convolución	2
Kernel size primera capa	5
Pooling size primera capa	3
Filtro primera capa	64
Kernel size segunda capa	3
Pooling size segunda capa	2
Filtro primera segunda	128

Cuadro 7: Elementos que optimizan la red neuronal convolucional.

se realiza un análisis de la salida de la red neuronal convolucional (Figura 36) y su curva ROC (Figura 37).

Analizando la salida de la red neuronal, se determina que es capaz de clasificar los eventos. Ambas distribuciones presentan máximos en sus etiquetas correspondientes. El tamaño de las colas de eventos son menores en comparación con las obtenidas en la ANN, por lo que nos encontramos ante un mejor clasificador.

Para la curva ROC, se observa que el área bajo ella es mayor en comparación con la ANN. Se obtiene un área de 0.7913 mientras que para la ANN fue de 0,6860. Para unos poderes de rechazo del 90 % y del 50 % los valores de la aceptación son del 28,1 % y del 96,4 % respectivamente. Ambas aceptancias son mayores para la CNN. Se produce un menor rechazo de eventos de retrocesos nucleares, por lo que se pierde menos información.

### 6.3. Clasificación en función de la posición reconstruida

Se procede a volver a buscar patrones en los eventos que han sido catalogados como “TP”, “FP”, “TN” y “FN”. Se debe esperar un comportamiento similar al que se obtiene para la ANN. Sólo se tendrán en cuenta los poderes de rechazo del 50 % y del 90 %. Para estos dos casos, se representan las distribuciones en la posición Z reconstruida para todas las 4 muestras de polvo juntas, como se observa en las Figuras 38 y 39.

Para un poder del rechazo del 50 % (Figura 38), predominan de retrocesos nucleares correctamente clasificados (true nuclear recoil events) frente a los erróneos (false dust events). De forma contraria, existe una mayor cantidad de eventos de polvo clasificados erróneamente (false nuclear events) frente a los correctos (true dust events). Se cumple el mismo comportamiento que para la ANN. Para este caso, se catalogan muy pocos eventos

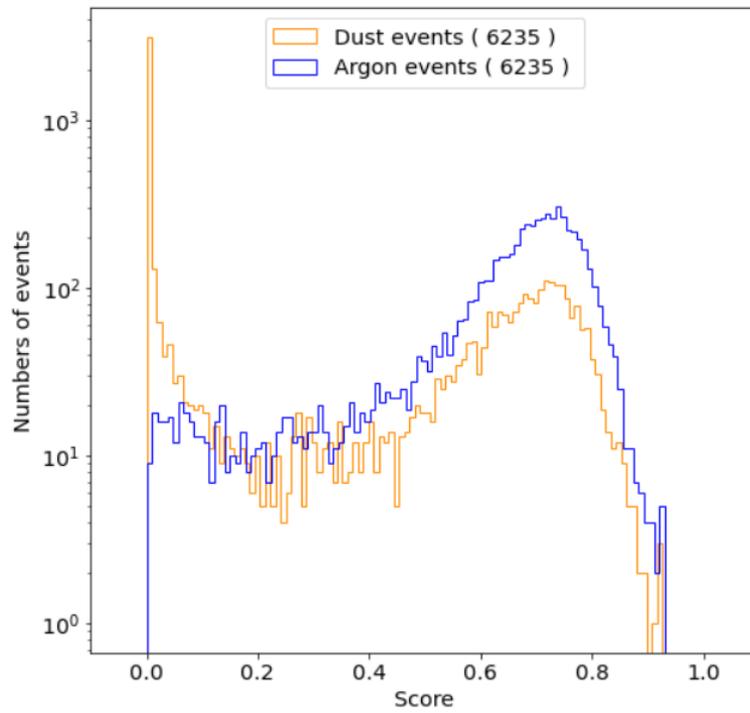


Figura 36: Distribución del resultado de la red CNN para las poblaciones de retrocesos nucleares y eventos de polvo de 25, 17, 10 y 5 micras que forman el conjunto de test tras aplicar SMOTE.

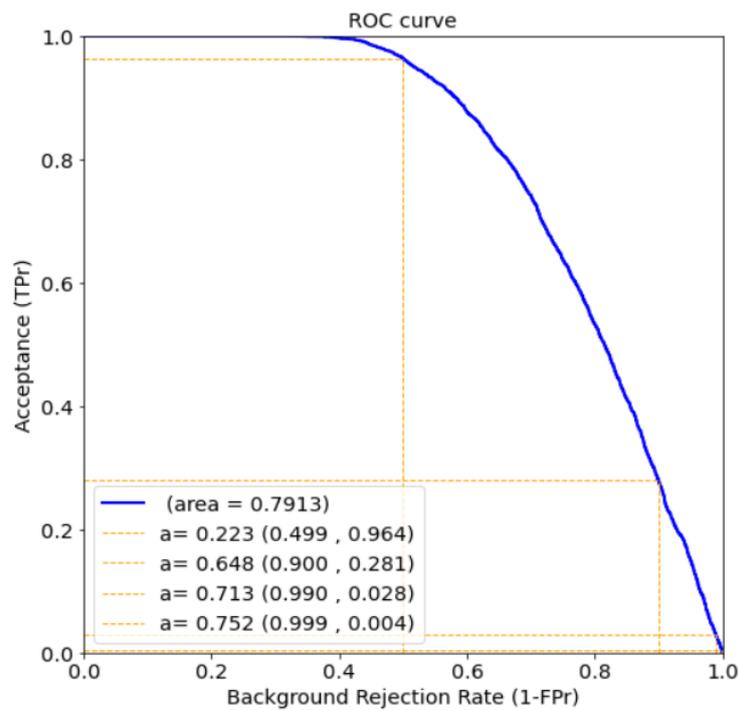


Figura 37: Curva ROC de la CNN tras aplicar SMOTE al conjunto de test de eventos de polvo de 5, 10, 17 y 25 micras.

como polvo para un umbral de  $u = 0,223$ . Este resultado implica que los eventos reales de polvo a la salida de la red toman valores no muy próximos a 0.

Cuando el poder de rechazo es del 90 % (Figura 39), sucede la situación contraria otra vez. Predominan los errores de retrocesos nucleares y los aciertos de eventos de polvo. En este caso, el umbral establecido es  $u = 0,648$ , situado más próximo a 1 que el caso anterior. Existirán más eventos de cola de retrocesos nucleares que serán erróneamente clasificados, aunque realiza una mejor clasificación de los eventos de polvo.

El comportamiento de la CNN a la hora de clasificar los eventos en función de la posición reconstruida es el mismo que para la ANN. Existe una mayor cantidad de eventos en el hemisferio superior del detector. En las partes superiores e inferiores del detector, debido a un menor volumen, no se detectan apenas eventos. El algoritmo no es capaz de distinguir ninguna propiedad. No se pueden extraer patrones característicos que distingan ambos tipos de eventos. Poseen las mismas distribuciones otra vez.

La red neuronal clasifica otra vez mejor los eventos sintéticos de polvo. Para los eventos originales de polvo, su clasificación es mejor para un poder de rechazo del 90 % frente al 50 %. Se clasifican correctamente el 74,9 % y el 0,45 % de los eventos originales para ambos poderes de rechazo respectivamente. Los valores de aceptación de retrocesos nucleares son los obtenidos otra vez por las curvas ROC. Debido a que la colaboración no exige un mínimo de poder de rechazo, el 90 % es aquel que mejor clasifica los eventos de polvo reales obteniendo una mayor aceptación de retrocesos nucleares.

#### 6.4. Clasificación en función de la energía

Una vez analizado el patrón de la posición reconstruida, se procede a analizar de nuevo el patrón de energía. De la misma forma, se estudian los poderes de rechazo del 50 % y del 90 %. Se representa las distribuciones de la energía para todas las 4 muestras de polvo (Figuras 40 y 41).

Se observa la misma distribución de eventos que para la ANN. Para un poder de rechazo del 50 % (Figura 40) predominan los aciertos de retrocesos nucleares y los errores de polvo. Se catalogan muy pocos eventos como polvo. Para un poder de rechazo del 90 % (Figura 41) predominan los aciertos de polvo y los errores de retrocesos nucleares. La justificación es igualmente el umbral obtenido.

Al igual que para la ANN, se puede determinar un patrón característico en la energía de los retrocesos nucleares. La red realiza esta asociación y la sobreexplota para clasificar los eventos. Es necesario realizar un trabajo futuro para desarrollar cortes que eviten

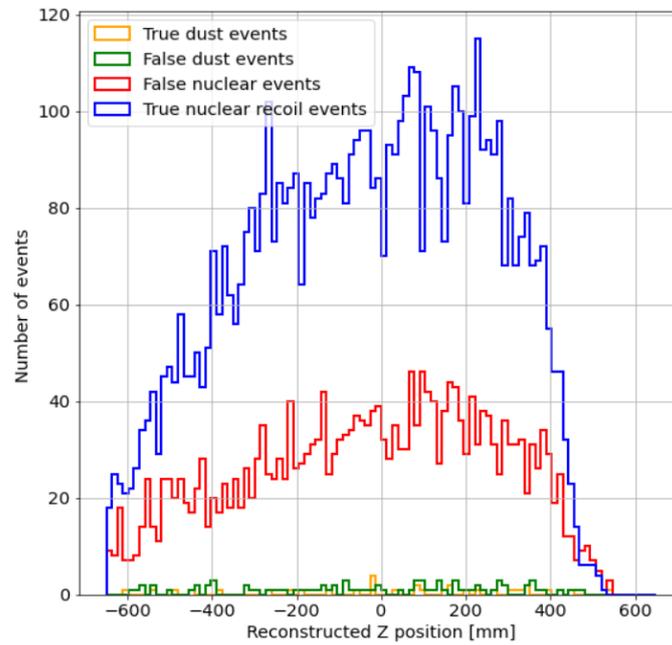


Figura 38: Histograma de posición reconstruida para un poder de rechazo del 50% sobre el conjunto datos de test original de todas las muestras de polvo empleando la CNN.

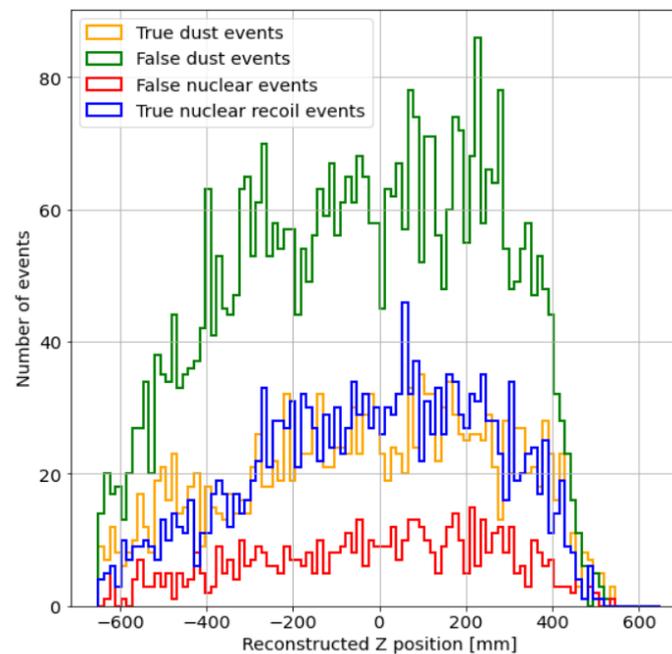


Figura 39: Histograma de posición reconstruida para un poder de rechazo del 90% sobre el conjunto datos de test original de todas las muestras de polvo empleando la CNN.

que la red actúe mediante este patrón.

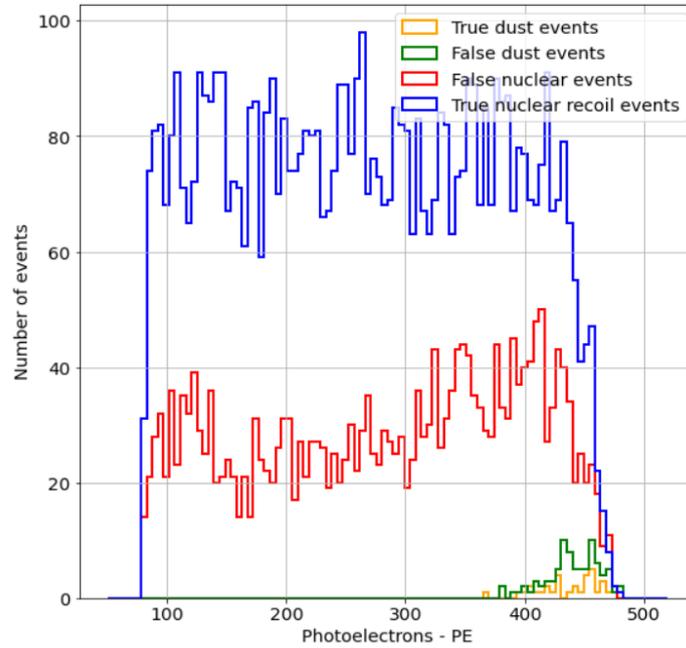


Figura 40: Histograma de energía para un poder de rechazo del 50% sobre el conjunto de datos de test original de todas las muestras de polvo en la red convolucional.

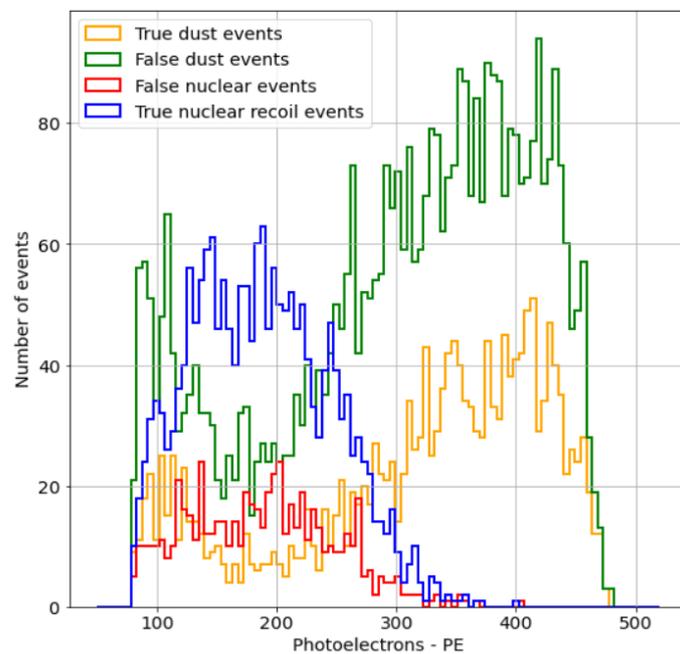


Figura 41: Histograma de energía para un poder de rechazo del 90% sobre el conjunto de datos de test original de todas las muestras de polvo en la red convolucional.

### 6.5. Validación sobre datos de $^{39}\text{Ar}$

En este apartado, mediante la red neuronal convolucional, se vuelve a realizar la prueba de verificación de los datos. Se debe esperar que la salida vuelva a ser la misma para eventos reales y simulados. El conjunto de datos es el mismo, 100874 eventos de datos reales del detector y 17484 eventos de datos simulados.

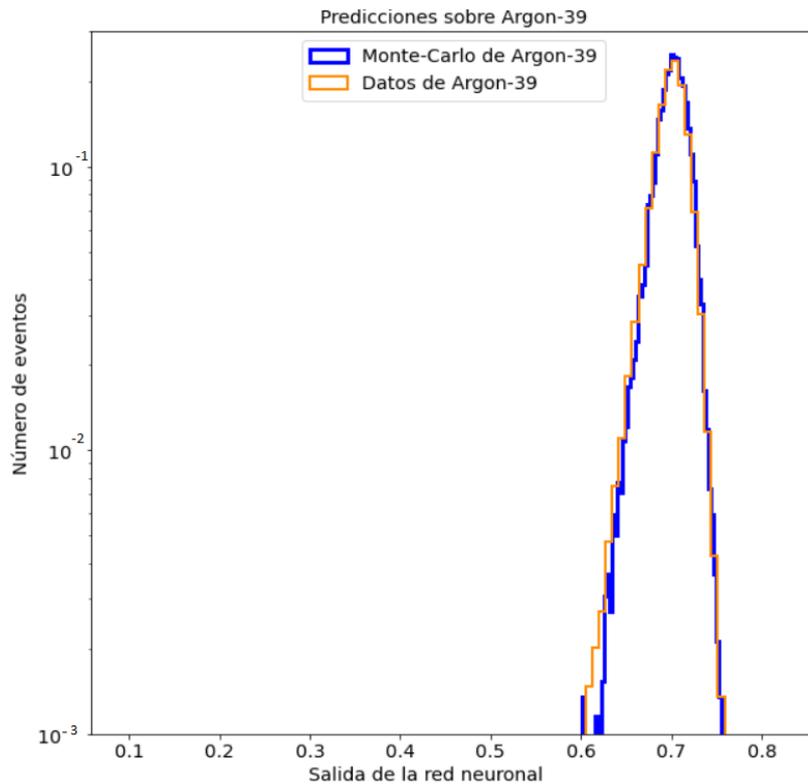


Figura 42: Aplicación de la mejor red tipo CNN sobre datos de retrocesos electrónicos experimentales y simulados, producidos por el decaimiento beta del  $^{39}\text{Ar}$ . Ambos subconjuntos están normalizados para poner énfasis en las distribuciones.

En la Figura 42 se observan las distribuciones normalizadas, tanto para los datos simulados y para los datos reales, aplicando la misma red neuronal tipo cnn una vez entrenada con las simulaciones. Se muestra que ambas distribuciones son muy similares otra vez entre ellas, lo que implica que la red no distingue entre realidad y simulación. Es posible implementar la red para analizar datos reales.

## 7. Conclusiones y trabajo futuro

En este trabajo se han presentado distintas soluciones para mejorar la clasificación, con redes neuronales, de diferentes tipos de eventos en el experimento de materia oscura DEAP-3600. El fondo más problemático estudiado ha sido la aparición de partículas de polvo dentro del volumen del detector, cuyas desintegraciones  $\alpha$  provocan señales muy similares a las que producen los WIMPs.

Tras realizar un estudio de los distintos eventos de retrocesos nucleares y de fondo de polvo proporcionados, la solución final propuesta ha sido emplear una red neuronal convolucional, la cual explota factores geométricos como la proximidad entre los fotomultiplicadores. Además, para poder realizar un entrenamiento efectivo de la red, ha sido necesario evitar el desbalanceamiento de los datos aplicando la técnica SMOTE sobre los eventos de muestras de polvo.

Después de un proceso de optimización, la red es capaz de obtener, para un poder de rechazo del 90 % una aceptación del 28,1 %. Finalmente, la red ha sido validada mediante eventos reales de  $^{39}\text{Ar}$ . No se ha observado ninguna diferencia entre ambas distribuciones de eventos.

La baja aceptación para poderes de rechazo mayores implica que se pierde una gran cantidad de los eventos de retrocesos nucleares, por lo que se pueden proponer desarrollos futuros que puedan mejorar los resultados. Existen varios trabajos enfocados a obtener una mayor estadística de datos para ver si la preparación de los datos usando SMOTE juega algún papel. Aproximadamente un 85 % de los eventos de polvo son producidos mediante SMOTE. También se propone una búsqueda para determinar si los eventos sintéticos poseen alguna relación con la dependencia en energía y poder aumentar la aceptación de retrocesos nucleares para mismos valores de poder de rechazo.

## Agradecimientos

Finalmente, me gustaría agradecer a mis directores, Miguel y Vicente. Gracias por todos los consejos y ayudas que me habéis aportado durante este año, tanto en lo profesional como en lo personal. También al grupo de Materia Oscura del CIEMAT y a la colaboración del experimento DEAP-3600 por brindarme esta gran oportunidad.

No me quiero olvidar de mi familia, que ha estado apoyándome día tras día. Tampoco de mis amigos y colegas, cuyos consejos con cerveza en mano me han animado a seguir adelante con este proyecto. Por último, recuerden, ¡en esta casa obedecemos las leyes de

la TERMODINÁMICA!

## 8. Referencias

- [1] N. Aghanim et al. *Planck 2018 Results. VI. Cosmological parameters*. A&A. arXiv:1807.06209 (2018)
- [2] M. Lozano, A. Casas. *La materia oscura. El elemento más misterioso del universo*. National Geograpic. RBA Editorial (2015).
- [3] F. Zwicky. *Helvetica Physica Acta*. 6: 110-127 (1933).
- [4] F. Zwicky. *Astrophysical Journal*. 86: 217 (1937).
- [5] K. Freese. *Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars*. Conference Proceeding for "Dark Matter and Dark Energy", 14, 7 (2008).
- [6] A. Abdulkhalaq. *Introduction to Black Holes*. M.SC. Gravity, Particle and Fields, 2110603 (2012).
- [7] S. Molina. *Búsqueda de materia oscura: situación actual y perspectivas de futuro*. Universidad Internacional de Valencia (2018).
- [8] M. Markevitch et al. *A direct empirical proof of the existence of dark matter*. ApJ 606, 819 (2006).
- [9] K. Griest and M. Kamionkowski. *Supersymmetric dark matter*. Phys. Rept. 333-334, 167-182 (2000).
- [10] F. Rota Font. *Extensiones del modelo estándar del universo primitivo nucleosíntesis primordial, axiones y materia oscura*. Universitat Autònoma de Barcelona (2006).
- [11] E. Armengaud et al. *Conceptual design of the International Axion Observatory (IAXO)*. Journal of Instrumentation 9.05 (2014).
- [12] JR. Minkel. *Experiment Nixes Fourth Neutrino*. Scientific American (2007).
- [13] M. Sáez, M. Mosquera and E. Civitarese. *Neutrinos estériles en nucleosíntesis primordial*. Asociación Argentina de Astronomía (2015).
- [14] A. Aguilar-Arevalo et al. *LSND Collaboration*. Phys. Rev. D 64, 112007 (2001).
- [15] M. Milgrom. *A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis*. ApJ, 270, 365-370. (1983).

- [16] M. Aguilar et al. *Towards Understanding the Origin of Cosmic-Ray Positrons*. Phys. Rev. Lett. 122, 0401102 (2019).
- [17] G. Ambrosini et al. *Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons*. Nature, 552, 63-66 (2017).
- [18] A.M. Sirunyan, A. Tumasyan et al. *Observation of  $t t H$  production*. J. High Energ. Phys. 3 160 (2018).
- [19] B. Montés Núñez. *Analysis of the first underground run and background studies of the Argon Dark Matter*. PhD, UCM (2016).
- [20] L. Baudis et al. *Dark matter search results from a one ton-year exposure of XENON1T*. Physical Review Letters, vol. 121, no 11, p. 111302 (2018).
- [21] P. Agnes et al. *Low-mass dark matter search with the DarkSide-50 experiment*. Physical review letters, 121(8), 081307 (2018).
- [22] R. Amaudruz et al. *Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB*. Physical Review D, 100(2), 022004 (2019)
- [23] J.D. Lewin and P.F. Smith. *Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil*. Astropart. Phys, 6, 87 (1996).
- [24] P. Cushman et al. *Snowmass CF1 summary: WIMP dark matter direct detection*. arXiv preprint arXiv:1310.8327 (2013).
- [25] C. Yaguna. *Isospin-violating dark matter in the light of recent data*. Physical Review D, 95, 05, 055015 (2017).
- [26] P.-A. Amaudruz et al. *Measurement of the scintillation time spectra and pulse-shape discrimination of low-energy beta and nuclear recoils in liquid argon with DEAP-1*. Astroparticle Physics 85, 1-23 (2016).
- [27] I. García. *Desarrollo de una red neuronal para rechazar eventos de fondo en el detector de materia oscura DEAP-3600*. UCM (2019).
- [28] R. Amaudruz et al. *Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB*. Physical Review D, 100(2), 022004 (2019)
- [29] J. Torres. *Deep Learning: Introducción práctica con Keras*. (2018). Versión online en <http://torres.ai>

- [30] I. Goodfellow, Y. Bengio and A. Courville. *Deep Learning*. MIT Press, (2016).
- [31] D. Matich. *Redes Neuronales: Conceptos básicos y aplicaciones*. Universidad Tecnológica Nacional, México, match2001redes (2001).
- [32] P. Larranaga, I. Inza, and A. Moujahid. *Tema 8. Redes neuronales*. U. del P. Vasco, 12, 17 (1997).
- [33] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. *Dropout: a simple way to prevent neural networks from overfitting*. The journal of machine learning research, 15, 1 (2014).
- [34] F. Chollet et al. *Keras*. <https://keras.io> (2015).
- [35] F. Pedregosa et al. *Scikit-learn: Machine learning in Python*. *the Journal of machine Learning research*. vol. 12, p. 2825-2830. (2011).
- [36] J. Monroy-de-Jesús, A. Guadalupe-Ramírez, JC Ambriz-Polo and E. López-González, *Algoritmo de aprendizaje eficiente para tratar el problema del desbalance de múltiples clases*. Tecnológico de Estudios Superiores de Jocotitlán, IOP Publishing (2020)
- [37] HA. Carnejo Calviño. *Métodos para la mejora de predicciones en clases desbalanceadas en el estudio de bajas de clientes (CHURN)*. USC (2017).