



Universidad de San Carlos de Guatemala
Escuela de Ciencias Físicas y Matemáticas
Departamento de Física

RECONSTRUCCIÓN DE LA ENERGÍA DE RAYOS CÓSMICOS DETECTADOS POR HAWC UTILIZANDO REDES NEURONALES

Daniel Alexander Alvarado Mérida

Asesorado por Dr. Tomás Capistrán Rojas y Dr. José Rodrigo Sacahuí Reyes

Guatemala, septiembre de 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

**RECONSTRUCCIÓN DE LA ENERGÍA DE RAYOS
CÓSMICOS DETECTADOS POR HAWC
UTILIZANDO REDES NEURONALES**

TRABAJO DE GRADUACIÓN
PRESENTADO A LA JEFATURA DEL
DEPARTAMENTO DE FÍSICA
POR

DANIEL ALEXANDER ALVARADO MÉRIDA
ASESORADO POR DR. TOMÁS CAPISTRÁN ROJAS Y DR. JOSÉ RODRIGO
SACAHUÍ REYES

AL CONFERÍRSELE EL TÍTULO DE
LICENCIADO EN FÍSICA APLICADA

GUATEMALA, SEPTIEMBRE DE 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS



CONSEJO DIRECTIVO INTERINO

Director	M.Sc. Jorge Marcelo Ixquiac Cabrera
Representante Docente	Arqta. Ana Verónica Carrera Vela
Representante Docente	M.A. Pedro Peláez Reyes
Representante de Egresados	Lic. Urías Amitaí Guzmán García
Representante de Estudiantes	Elvis Enrique Ramírez Mérida
Representante de Estudiantes	Oscar Eduardo García Orantes
Secretario	M.Sc. Freddy Estuardo Rodríguez Quezada

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

Director	M.Sc. Jorge Marcelo Ixquiac Cabrera
Examinador	Dra. Ana Beatriz Cosenza Muralles
Examinador	M.Sc. Edgar Anibal Cifuentes Anleu
Examinador	M.Sc. Osmar Obdulio Hernández Aguilar
Secretario	M.Sc. Freddy Estuardo Rodríguez Quezada

Ref. D.DTG. 011-2024
Guatemala 24 de septiembre de 2024

El Director de la Escuela de Ciencias Físicas y Matemáticas de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del jefe de la Licenciatura en Física Aplicada, al trabajo de graduación titulado: " **RECONSTRUCCIÓN DE LA ENERGÍA DE RAYOS CÓSMICOS DETECTADOS POR HAWC UTILIZANDO REDES NEURONALES** ", presentado por el estudiante universitario, Daniel Alexander Alvarado Mérida, autoriza la impresión del mismo.

IMPRÍMASE.

"ID Y ENSEÑAD A TODOS"



M.Sc. Jorge Marcelo Ixquiac Cabrera
Director

AGRADECIMIENTOS

A Las Amebas, por estar desde el inicio y crecer conmigo durante todo este tiempo.

A Sarce y Chepi por todo el aprendizaje juntos.

Al Grupo Euclideano, por todas las tardes de pandemia estudiando.

Al Team Spaces por escuchar mis quejas.

A mis asesores, Rodrigo y Tomás por la infinita paciencia que tuvieron durante este proyecto. Asimismo agradezco a Ibrahim Torres por compartir su conocimiento desinteresadamente y por levantarme el ánimo después de las reuniones.

Gracias por recordarme de lo que soy capaz y por escuchar mis quejas todo este tiempo.

I am. All of us.

AGRADECIMIENTOS A PROGRAMAS

Este proyecto fue realizado gracias al apoyo de los proyecto PAPIIT IG101323 y proyecto de laboratorios nacionales de CONAHCyT LNC-2023-117.

Agradezco al programa de becas en la UNAM de DGAPA por el apoyo brindado para el desarrollo de este proyecto.

DEDICATORIA

A mis padres, por enseñarme a trabajar duro. Por su infinito apoyo y respaldo.

A mis hermanas, por alegrarme la vida.

A Mauricio, Emilio, Ñoqui (Q.D.E.P.) y Carmelo (Q.D.E.P.).

ÍNDICE GENERAL

ÍNDICE DE FIGURAS	IV
ÍNDICE DE TABLAS	V
LISTA DE SÍMBOLOS	VII
OBJETIVOS	IX
INTRODUCCIÓN	XI
1. CONCEPTOS PRELIMINARES	1
1.1. Rayos cósmicos	1
1.1.1. Contexto histórico	1
1.2. Cascadas atmosféricas extendidas	2
1.2.1. Cascadas electromagnéticas	2
1.2.2. Cascadas hadrónicas	3
1.2.3. Características importantes de las EAS	5
1.3. Espectro de energía de los rayos cósmicos	5
1.4. Aprendizaje automático	6
1.4.1. Tipos de aprendizaje	7
1.4.1.1. Aprendizaje supervisado	7
1.4.1.2. Aprendizaje por refuerzo	7
1.4.1.3. Aprendizaje no supervisado	7
1.4.2. Redes neuronales	7
1.4.2.1. El perceptrón	8
1.4.2.2. El perceptrón multicapa (MLP)	8
1.4.2.3. Entrenamiento de redes neuronales	9
2. EL OBSERVATORIO HAWC	13
2.1. Características generales	13

2.2. Instrumentación y técnica	14
2.2.1. Efecto Cherenkov	14
2.2.2. Tubos fotomultiplicadores (PMTs)	14
2.2.3. Detectores Cherenkov en agua (WCD)	15
2.3. Software en HAWC	16
2.3.1. CORSIKA	16
2.3.2. HAWCSim	16
2.3.3. HAWCSim-Reco	16
3. Metodología	17
3.1. Entrenamiento de modelos de redes neuronales	17
3.1.1. Selección de datos de entrenamiento	17
3.1.2. Arquitectura de los modelos entrenados	18
3.1.2.1. Parámetros de entrada para el entrenamiento de modelos	18
3.1.3. Modelos de redes neuronales entrenados	19
3.2. Cortes de calidad y cortes de selección de eventos	20
4. RESULTADOS	23
4.1. Comparación entre modelos	23
4.1.1. Comparación entre modelos de redes neuronales	23
4.1.2. Comparación del tercer modelo contra el estimador actual de HAWC	24
4.2. Pruebas en eventos iniciados por núcleos de hierro	26
4.3. Discusión y conclusiones	27
BIBLIOGRAFÍA	29

ÍNDICE DE FIGURAS

1.1.	Representación simple de una cascada electromagnética. Tomado de [16].	3
1.2.	Representación de una cascada hadrónica iniciada por un protón. Tomado de [16].	4
1.3.	Espectro diferencial de energía de los rayos cósmicos primarios. Se señalan los quiebres <i>knee</i> y <i>ankle</i> , conocidos en español como rodilla y tobillo, respectivamente. La región entre 10^6 eV y la rodilla se aproximan a un ley de potencia inversa de índice espectral $\gamma_1 = 2.68$, mientras de la rodilla al tobillo se aproxima a $\gamma_2 = 3.15$. Tomado de [15].	6
1.4.	Representación gráfica del MLP, tomado de [10].	10
1.5.	Representación de la función de error como una superficie en el espacio de pesos, tomado de [10].	11
1.6.	Representación del procedimiento de back propagation en una red neuronal. Tomado de [10].	12
2.1.	Fotografía del observatorio HAWC. Recuperado de: https://www.hawc-observatory.org/	13
2.2.	Frente de onda de la radiación cherenkov en un medio de índice de refracción n . Tomado de [16].	14
2.3.	A la izquierda: esquema de un PMT. A la derecha: vista física de un PMT utilizado en HAWC. Imágenes tomadas de [16].	15
2.4.	Ilustración de un detector Cherenkov usado en HAWC. Imagen tomada del sitio privado de HAWC.	15

3.1. Arquitectura de las redes neuronales entrenadas. La arquitectura se define con el número de neuronas en cada capa. En este caso, observamos la arquitectura 14:10:10:1, donde cada círculo blanco representa una neurona y estas están unidas por medio de flechas que representan pesos. Fuente: elaboración propia con TMVA.	19
4.1. Se muestra un histograma 2D de la energía reconstruida vs la energía verdadera de los tres modelos entrenados en este trabajo (a,b y c, respectivamente). Se muestra la identidad como una línea negra en la diagonal. Fuente: elaboración propia.	24
4.2. Histograma normalizado de $\Delta \text{Log}(E)$ en el intervalo (3.50-3.75) de $\log(E^{\text{True}}/GeV)$ y un ajuste gaussiano en rojo.	25
4.3. Histogramas 2D de la energía reconstruida por el estimador oficial de HAWC (a) y por el mejor modelo de redes neuronales (b). Fuente: elaboración propia.	26
4.4. Distribución del bias del estimador actual de HAWC (Likelihood) y el tercer modelo entrenado en este trabajo (NeuralNetwork). Fuente: elaboración propia.	26
4.5. Se muestra un histograma 2D ilustrando la energía verdadera contra la energía reconstruida por el estimador oficial de HAWC (a), y la energía reconstruida por el tercer modelo de redes neuronales (b) utilizando EAS inducidas por núcleos de hierro. Fuente: elaboración propia.	27
4.6. Comparación del bias de la energía reconstruida por el estimador actual de HAWC y el tercer modelo de redes neuronales para eventos iniciados por núcleos de hierro. Fuente: Elaboración propia.	27

ÍNDICE DE TABLAS

3.1. Parámetros de entrada utilizados y las información que contienen sobre las cascadas atmosféricas extensas. Fuente: elaboración propia.	19
3.2. Resumen de los modelos entrenados, el número de redes neuronales usadas, el detalle de los bins de <code>fhit</code> de cada red neuronal y si se usaron datos de núcleos encontrados dentro o fuera de HAWC.	20

LISTA DE SÍMBOLOS

Símbolo	Significado
$:=$	es definido por
\cong	es isomorfo a
e	electrón
e^+	positrón
γ	fotón
π^+	pión positivo
π^-	pión negativo
π	pión neutro
μ^-	muón
μ^+	antimuón
ν_μ	neutrino muónico
$\bar{\nu}_\mu$	antineutrino muónico
eV	electronvoltio
∇	operador diferencial nabra en n-dimensiones

OBJETIVOS

General

Entrenar una red neuronal para estimar la energía de los eventos hadrónicos detectados por el observatorio HAWC y cuantificar su rendimiento.

Específicos

1. Entrenar, evaluar y comparar diferentes arquitecturas de redes neuronales utilizando simulaciones Monte Carlo de cascadas atmosféricas iniciadas por protones.
2. Seleccionar el mejor modelo basado en su desempeño al reconstruir eventos iniciados por protones.
3. Comparar el desempeño del mejor modelo contra el actual estimador de energías de rayos cósmicos utilizado en HAWC.
4. Evaluar la robustez y generalidad del modelo seleccionado empleando núcleos de hierro.

INTRODUCCIÓN

Los rayos cósmicos son partículas que provienen del cosmos, estos se caracterizan por tener carga eléctrica y por viajar a velocidades cerca a la velocidad de la luz. Están compuestos por un 90 % de núcleos de hidrógeno (comúnmente llamados protones), un 10 % partículas alfa y un pequeño resto núcleos pesados [12]. Estos producen cascadas atmosféricas extensas (EAS, por sus siglas en inglés) al interactuar por medio de la fuerza nuclear con núcleos de la atmósfera [15]. Por otro lado, varios experimentos han recopilado información sobre el espectro de energía de los rayos cósmicos incidentes en la Tierra. Este está descrito por una ley de potencias de índice $\gamma \approx -2.7$ y tiene dos quiebres que forman una región con índice $\gamma \approx -3.15$. El primer quiebre se ubica alrededor de 3×10^{15} eV (llamado rodilla) y otro alrededor de 4×10^{18} eV (llamado tobillo) [12]. El espectro de energía es de interés para inferir propiedades físicas acerca de las fuentes de rayos cósmicos, los lugares del universo que los aceleran y los ambientes que atraviesan tras su llegada a la Tierra [16].

Existen diferentes formas de detectar rayos cósmicos, por ejemplo, con detectores satelitales como la misión Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) que tuvo fin en 2016; detectores a bordo de globos estratosféricos como el Cosmic Ray Energetics and Mass (CREAM); y observatorios en tierra como el observatorio Pierre-Auger y el High Altitude Water Cherenkov Experiment (HAWC). En particular, el observatorio HAWC está ubicado en las faldas del Volcán Sierra Negra en Puebla, México. Este consiste en un arreglo de 300 detectores Cherenkov en agua desplegados en un área de $22,000\text{m}^2$ [2]. Si bien HAWC está optimizado para detectar rayos gamma, su ubicación y técnica de detección permite recolectar información sobre rayos cósmicos primarios desde unos cientos de TeV hasta algunos PeV y realizar contribuciones en este ámbito, por ejemplo el estudio del espectro de energía de eventos hadrónicos [6].

Las investigaciones sobre rayos cósmicos en HAWC han empleado un estimador basado en métodos de máxima verosimilitud desarrollado en [16]. Sin embargo, este modelo presenta inconsistencias en su desempeño tras las recientes actualizaciones

en el software análisis de datos de HAWC para mejorar la reconstrucción de las EAS, por lo que se desea hacer una actualización en el estimador de energía o desarrollar un nuevo estimador de energía.

En este trabajo utilizamos técnicas de ciencia de datos que nos permiten desarrollar modelos para la reconstrucción de la energía de las EAS observadas por HAWC. Específicamente, se entrenaron redes neuronales usando datos de simulaciones de EAS iniciadas por protones (debido a su abundancia) hechas con el software oficial de HAWC. Estas simulaciones contienen parámetros con los cuales se reconstruyen las características más relevantes de las EAS y la respuesta de los detectores ante estos eventos. Luego de explorar algunas arquitecturas de redes neuronales, se evalúa la reconstrucción de la energía de estas y se selecciona la de mejor desempeño para ser comparada contra el estimador actual de HAWC. Finalmente, se evalúa la robustez y generalidad del modelo siendo sometido a reconstrucciones de eventos iniciados por otros hadrones

1. CONCEPTOS PRELIMINARES

1.1. Rayos cósmicos

1.1.1. Contexto histórico

En 1912, el físico austriaco Victor Hess realizó un experimento mediante un globo aerostático para medir la ionización atmosférica a distintas alturas. Hess notó que la cantidad de radiación ionizante aumentaba a medida que la altura aumentaba, lo que sugirió la existencia de una fuente de radiación externa a la Tierra. Este experimento se llevó a cabo a una altitud superior a 5.3 kilómetros, donde Hess pudo observar una cantidad significativamente mayor de radiación ionizante en comparación con las mediciones realizadas en la superficie terrestre.

En un inicio se pensó que los rayos cósmicos eran radiación gamma debido a su gran poder de penetración. Millikan presentó la hipótesis de que los rayos gamma se generaban a partir de la formación de núcleos de helio por parte de protones y electrones en el espacio interestelar.

Entre 1927 y 1928, el físico holandés Clay llevó a cabo un experimento que consistió en la medición de la variación de la intensidad en función de la latitud geomagnética. A través de sus mediciones, Clay descubrió que la ionización aumentaba a medida que se incrementaba la latitud, lo que permitió demostrar que los rayos cósmicos interactuaban con el campo geomagnético y, por tanto, están constituidos por partículas cargadas eléctricamente. [12].

Tras al menos 100 años de trabajo e investigación en rayos cósmicos, se ha determinado que estos son partículas cargadas que viajan a velocidades relativistas. Debido a su carga eléctrica, estas son deflectadas por campos magnéticos presentes en el espacio interestelar e intergaláctico. Además, se ha logrado determinar que los rayos cósmicos están compuestos por un 90 % de protones, un 10 % partículas alfa y un 1 % núcleos pesados.

1.2. Cascadas atmosféricas extendidas

Las cascadas atmosféricas extendidas (o EAS, por sus siglas en inglés) son cascadas de partículas que se producen a partir de rayos gamma o rayos cósmicos al interactuar con moléculas en la atmósfera terrestre. Los rayos gamma producen cascadas electromagnéticas, mientras que los rayos cósmicos producen cascadas hadrónicas. Las EAS se caracterizan por formar un disco de partículas delgado y de radio amplio, este disco se propaga a velocidades cercanas a la velocidad de la luz a lo largo del vector de momentum de la partícula primaria [15]. Hoy en día, las EAS son detectadas principalmente por arreglos de detectores Cherenkov de agua o aire, los cuales son capaces de reconstruir propiedades como la energía, ángulo de incidencia, entre otras propiedades. El análisis de las EAS ayuda a comprender la interacción de los rayos cósmicos con la atmósfera terrestre y su origen.

1.2.1. Cascadas electromagnéticas

Las cascadas electromagnéticas son iniciadas por rayos gamma incidentes en la Tierra. Constan principalmente de fotones, electrones y positrones. Los procesos predominantes en estas cascadas es la producción de pares electrón-positrón, descrita por:

$$\gamma \rightarrow e^- + e^+,$$

y radiación de frenado provocado por la desaceleración de una partícula cargada y de baja masa bajo la influencia del campo eléctrico de otra partícula cargada (bremsstrahlung). Este proceso está descrito por:

$$e^\pm + A \rightarrow e^\pm + \gamma + A.$$

Heitler [18] propone un modelo simple para describir el desarrollo de las EAS inducidas por rayos gamma (ver figura 1.1). En este modelo se asume que las interacciones de las partículas secundarias se dan entre longitudes de interacción λ , de forma que después de n longitudes λ se tienen $N = 2^n$ partículas secundarias. Otra consideración importante es que la energía primaria E_0 se distribuye equitativamente entre todas las partículas secundarias, es decir, la energía de la n -ésima partícula está dada por

$$E_n = E_0/2^n. \tag{1.1}$$

La cascada continúa creciendo hasta que las partículas tienen una energía crítica E_c , la cual depende del medio en que se muevan las partículas. A partir de este momento, ya no se producen partículas secundarias y se ha llegado al máximo de interacciones, habiéndose propagado una longitud $X_{max} = \lambda \log_2(E_0/E_c)$ y con número de partículas igual a $N_{max} = E_0/E_c$.

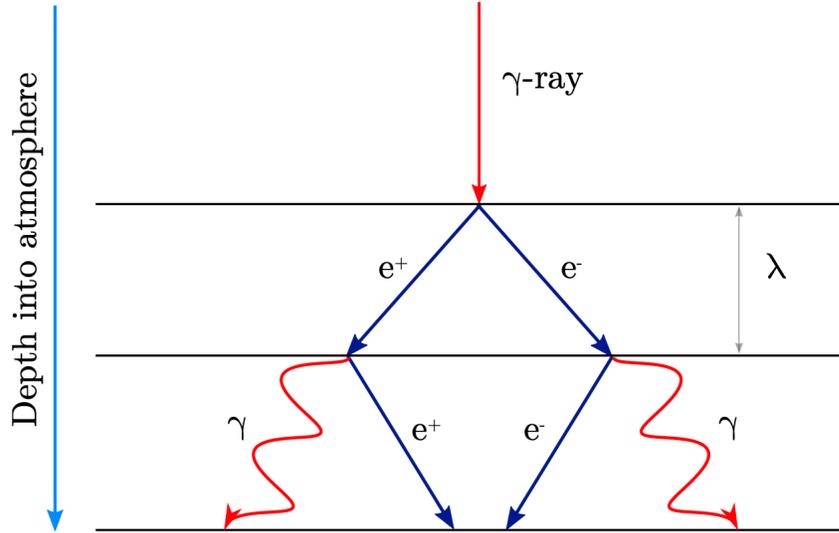


Figura 1.1. Representación simple de una cascada electromagnética. Tomado de [16].

1.2.2. Cascadas hadrónicas

Las partículas hadrónicas incidentes en la Tierra inducen EAS que se desarrollan de manera similar a las cascadas electromagnéticas. Sin embargo, a diferencia de las cascadas electromagnéticas, estas están gobernadas por procesos nucleares. Este tipo de EAS se dan por una interacción nuclear entre la partícula incidente (conocida como primario) y los núcleos presentes en la atmósfera. Esta interacción da lugar a la generación de partículas secundarias, que a su vez colisionan con otros núcleos produciendo varias generaciones más de partículas secundarias (ver figura 1.2). La cascada resultante está determinada principalmente por la energía del primario, el tipo de partícula, el ángulo de incidencia y la altura a la que se dio la primera interacción. Estas colisiones son predominantemente hadrónicas y la forman cascadas que se propagan longitudinalmente en dirección del momentum inicial del primario, sin embargo, también se propagan lateralmente debido al momentum transversal y procesos de dispersión de los secundarios [15].

Las cascadas iniciadas por hadrones están compuestas por una parte electro-

magnética (γ, e^\pm) y una parte hadrónica donde sus componentes principales son piones neutros y cargados (π^0, π^\pm). Los piones neutros contribuyen a la parte electromagnética vía

$$\begin{aligned} \pi^0 &\rightarrow 2\gamma \\ &\downarrow \\ \gamma &\rightarrow e^+ + e^- \end{aligned}$$

Mientras que los piones cargados, al colisionar con núcleos de la atmósfera decaen en muones vía

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu, \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu. \end{aligned}$$

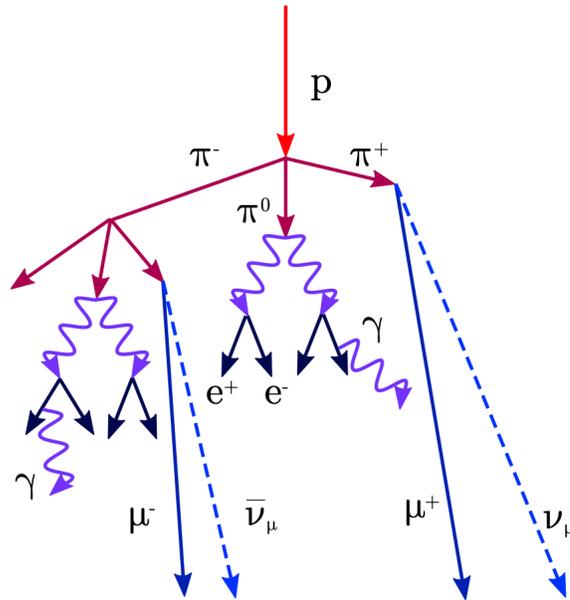


Figura 1.2. Representación de una cascada hadrónica iniciada por un protón. Tomado de [16].

El desarrollo longitudinal y lateral de una cascada depende principalmente de la energía del primario que la inició. En eventos de baja energía, la cascada alcanza su máximo desarrollo arriba en la atmósfera y se extingue con el aumento de la profundidad atmosférica, dejando solo los componentes de muones y neutrinos para alcanzar la superficie. Mientras que en eventos de energía extremadamente alta, la cascada puede alcanzar su máximo desarrollo cerca del nivel del mar.

De manera similar, la energía del primario juega un papel importante en el número de partículas secundarias. El número de partículas secundarias que son producidas

en una cascada a un nivel particular en la atmósfera es llamado “el tamaño de la cascada”. El tamaño de la cascada también está determinado por el ángulo de incidencia del primario así como la altura en la que se dio la primera interacción.

1.2.3. Características importantes de las EAS

Las EAS son fenómenos físicos complejos que involucran numerosas interacciones entre partículas. Sin embargo, podemos estudiarlas como un todo y definir algunas de sus características más importantes:

- El tamaño, definido en la Sección 1.2.2.
- La profundidad de máximo desarrollo, se define el lugar de la atmósfera donde se dio el número máximo de partículas secundarias. Es función de la energía del primario.
- El núcleo, se refiere al centro de distribución de carga.
- El eje se refiere a la proyección del momentum lineal del primario.
- El plano, es un disco formado por partículas secundarias que van al frente de la cascada.

1.3. Espectro de energía de los rayos cósmicos

El espectro diferencial de energía de los rayos cósmicos se suele expresar en términos del flujo. Este está definido como el número de partículas primarias N con energía E que llegan a la Tierra por unidad de área A , ángulo sólido Ω , tiempo t y energía E :

$$\mathcal{F} = \frac{d^4N(E)}{dAd\Omega dt dE}. \quad (1.2)$$

El flujo observado está definido por ley de potencia de índice espectral $\gamma_1 \approx -2.7$. Se han observado dos quiebres: en 3×10^{15} eV y 4×10^{18} eV, los cuales son llamados rodilla y tobillo, respectivamente (ver figura 1.3). Estos quiebres llevan consigo un cambio en el índice espectral, siendo ligeramente más abrupto desde la rodilla hasta el tobillo, con un índice espectral $\gamma_2 \approx 3.15$. Se especula que el primer quiebre se deba a una característica de la fuente espectral o que sea causado por un cambio en la composición de la masa de los primarios. Por otro lado, se especula que

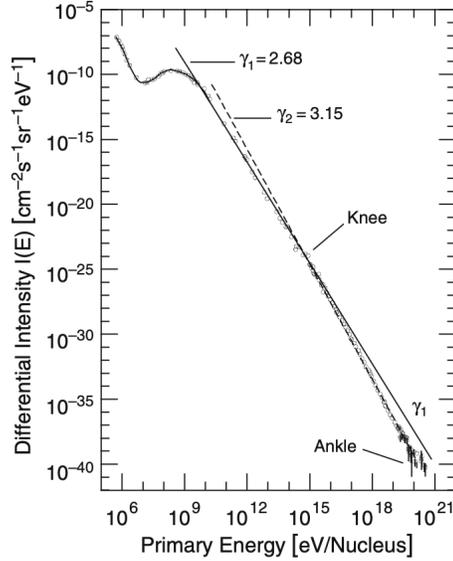


Figura 1.3. Espectro diferencial de energía de los rayos cósmicos primarios. Se señalan los quiebres *knee* y *ankle*, conocidos en español como rodilla y tobillo, respectivamente. La región entre 10^6 eV y la rodilla se aproximan a un ley de potencia inversa de índice espectral $\gamma_1 = 2.68$, mientras de la rodilla al tobillo se aproxima a $\gamma_2 = 3.15$. Tomado de [15].

el segundo quiebre se deba a contribuciones extragalácticas arriba de los 10^{10} GeV [15].

La naturaleza de este flujo lleva consigo información sobre las propiedades físicas de los lugares que aceleran a los rayo cósmicos y los medios que atraviesan, puesto que al estar cargados, sus trayectorias son modificadas por campos magnéticos presentes en el espacio [16].

1.4. Aprendizaje automático

El aprendizaje automático o machine learning es un conjunto de técnicas que hacen posible que una computadora aprenda a realizar una tarea. Es conveniente resaltar que existe una diferencia entre entrenar una computadora y programar una computadora. Entrenar implica utilizar conjuntos de datos y métodos estadísticos, mientras programar es establecer un conjunto de reglas duras y pasos detallados para resolver un problema.

1.4.1. Tipos de aprendizaje

Existen diferentes tipos de aprendizaje automático. Estos difieren en la forma en la que las computadoras son expuestas a los datos para su entrenamiento. Los tipos de aprendizaje más populares son el aprendizaje supervisado, por refuerzo y no supervisado.

1.4.1.1. Aprendizaje supervisado

El aprendizaje automático supervisado se caracteriza por utilizar un conjunto de datos de entrenamiento $\{(x_1, y_1^{true}), (x_2, y_2^{true}), \dots, (x_n, y_n^{true})\}$ que incluye información explícita sobre la salida correcta y_i^{true} para cada entrada x_i [3].

1.4.1.2. Aprendizaje por refuerzo

El aprendizaje por refuerzo se diferencia del aprendizaje supervisado en la información brindada en los datos de entrenamiento. Mientras en el aprendizaje supervisado se tienen datos de la forma (x_i, y_i^{true}) [3], los datos del aprendizaje por refuerzo contienen una entrada, una salida y una medida que cuantifica la calidad de la salida, es decir

(entrada, salida, calificación de la salida)

1.4.1.3. Aprendizaje no supervisado

En el aprendizaje no supervisado se tiene simplemente un conjunto de entradas $\{x_1, x_2, \dots, x_n\}$ y no se tiene información alguna sobre las salidas. Este tipo de aprendizaje puede verse como la tarea de encontrar patrones y estructuras en los datos de entrenamiento. Por ejemplo, podemos pensar en el problema de aprender un idioma nuevo basándonos solamente en la fonética. A pesar de no tener suficiente información para aprender de manera fluida un idioma, eventualmente, habremos aprendido varias características como la entonación del idioma, el sonido de las vocales y algunas consonantes, etc. [3]

1.4.2. Redes neuronales

Las redes neuronales son un sistema con unidades que intentan imitar a una neurona del cerebro humano. Estrictamente, son un mapeo f tal que $f : \mathcal{X} \rightarrow \mathcal{Y}$,

donde \mathcal{X} es el espacio de parámetros de entrada y \mathcal{Y} es el espacio de decisiones o predicciones hechas por el mapeo f .

1.4.2.1. El perceptrón

Antes de introducir las redes neuronales es conveniente revisar un modelo simple llamado perceptrón. Este es un modelo de clasificación binaria que toma $\{x_1, x_2, \dots, x_n\}$ entradas a las cuales se les asignan diferentes pesos $\{w_1, w_2, \dots, w_n\}$, dependiendo de la ponderación que nuestro modelo considere apropiada [3]. La salida de este modelo es el resultado de la suma ponderada $\sum_i w_i x_i$. Podemos definir una decisión con base en el resultado de esta salida y un umbral definido por nosotros como $-b$. Es decir:

$$\begin{aligned} \text{Decisión 1 si: } & \sum_{i=1}^n w_i x_i > -b, \\ \text{decisión 2 si: } & \sum_{i=1}^n w_i x_i < -b. \end{aligned}$$

Esta fórmula puede escribirse de la forma:

$$h(x) = \text{Sign} \left(\left(\sum_{i=1}^n w_i x_i \right) + b \right).$$

Donde la función $\text{sign}(x)$ puede tomar los valores $x = 1$ y $x = -1$, correspondiendo con la decisión 1 y 2. El término b es un umbral que determina si el algoritmo tomará la decisión 1 o 2, al parámetro b se conoce como bias u offset. El término x_0 se puede definir $x_0 = 1$ de forma que $b = w_0 x_0 = w_0$. Así, obtenemos la expresión simplificada

$$h(x) = \text{Sign} \left(\sum_{i=0}^n w_i x_i \right). \quad (1.3)$$

1.4.2.2. El perceptrón multicapa (MLP)

El multilayer perceptron (MLP por sus siglas en inglés) o perceptrón multicapa es un modelo de aprendizaje que busca mejorar al perceptrón, utilizando múltiples unidades basadas en el perceptrón (llamadas neuronas) que se organizan por capas. Si bien existe un parecido o una relación entre estos dos algoritmos, un MLP es una estructura más compleja y permite resolver problemas más complicados que las clasificaciones binarias.

Como mencionamos anteriormente, los MLP están conformados por neuronas

organizadas por capas. Para construir el modelo, tomemos la capa de entrada con n variables x_i de entrada y tomemos M combinaciones lineales a_j tales que

$$a_j = \sum_{i=0}^n w_{ji}^{(1)} x_i,$$

donde $j = 1, \dots, M$ y el superíndice de $w_{ji}^{(1)}$ indica que los parámetros pertenecen a la capa 1 de la red neuronal. De manera similar al perceptrón, los elementos w_{ji} con $i \neq 0$ son llamados pesos, mientras que los elementos $w_{j0}^{(1)}$ son los offsets. Por otro lado, las cantidades a_j son llamadas activaciones, cada una de las M activaciones deben ser transformadas por medio de una función diferenciable y no lineal $h(\cdot)$, de forma que definimos z_j como

$$z_j = h(a_j), \tag{1.4}$$

llamadas unidades escondidas. Ahora hacemos M combinaciones lineales de z_j , de forma que para la segunda capa tenemos las activaciones

$$a_k = \sum_{j=0}^M w_{kj}^{(2)} z_j, \tag{1.5}$$

donde $k = 1, \dots, K$ es el número de salidas. Esta salida a_k debe ser sometida también a una función de activación $\sigma(\cdot)$ tal que $y_k = \sigma(a_k)$, con y_k la salida de nuestra red neuronal. Al unir las piezas de nuestra red, tenemos que

$$y_k(\mathbf{w}, \mathbf{x}) = \sigma \left(\sum_{j=0}^M w_{kj}^{(2)} h \left(\sum_{i=0}^n w_{ji}^{(1)} x_i \right) \right). \tag{1.6}$$

Podemos ver una representación esquemática del MLP en la figura 1.4. [10].

1.4.2.3. Entrenamiento de redes neuronales

El proceso de entrenamiento de las redes neuronales consiste en encontrar los pesos adecuados \mathbf{w} que minimicen la función de error definida por un conjunto de vectores de entrada $\{\mathbf{x}_n\}$ y un conjunto correspondiente de vectores $\{\mathbf{t}_n\}$ los cuales modelan las salidas verdaderas de cada \mathbf{x}_n , esta función tiene la forma

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{y}(\mathbf{x}_n, \mathbf{w}) - \mathbf{t}_n\|^2. \tag{1.7}$$

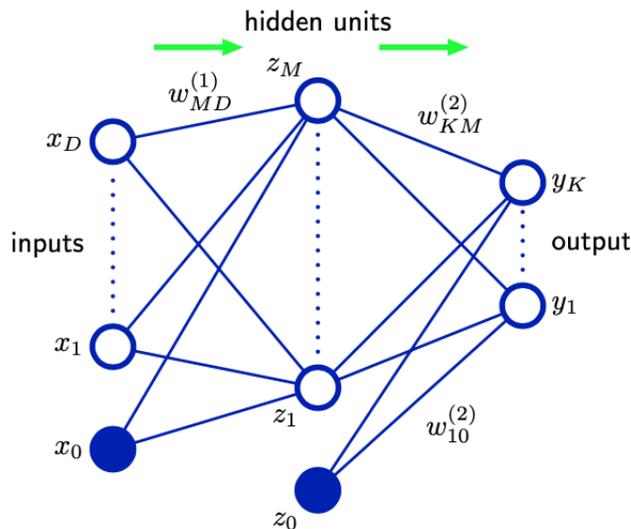


Figura 1.4. Representación gráfica del MLP, tomado de [10].

Dado que $E(\mathbf{w})$ es una función suave y continua de \mathbf{w} , el error disminuirá según el gradiente $-\nabla E(\mathbf{w})$ (ver figura 1.5). Los mínimos de esta función se encontrarán en lugares donde su gradiente sea cero, es decir, en

$$\nabla E(\mathbf{w}) = 0. \quad (1.8)$$

Esta optimización se hace con métodos numéricos dada la incerteza de la existencia de soluciones analíticas. En general, para que una red neuronal se desempeñe correctamente, no es necesario encontrar el mínimo global de la función de error. Encontrando suficientes mínimos locales y compararlos nos brinda una solución suficientemente buena [10]. El método para encontrar los pesos w_i es llamado back propagation, en el cual se toman derivadas parciales de la función de error con el propósito de evaluar el gradiente de la función de error de forma eficiente. Supongamos que nuestra función de error se puede expresar como

$$E(\mathbf{w}) = \sum_{n=1}^N E_n(\mathbf{w}). \quad (1.9)$$

Además, como vimos anteriormente, cada unidad de la red neuronal tiene la forma

$$a_j = \sum_i w_{ji} z_i, \quad (1.10)$$

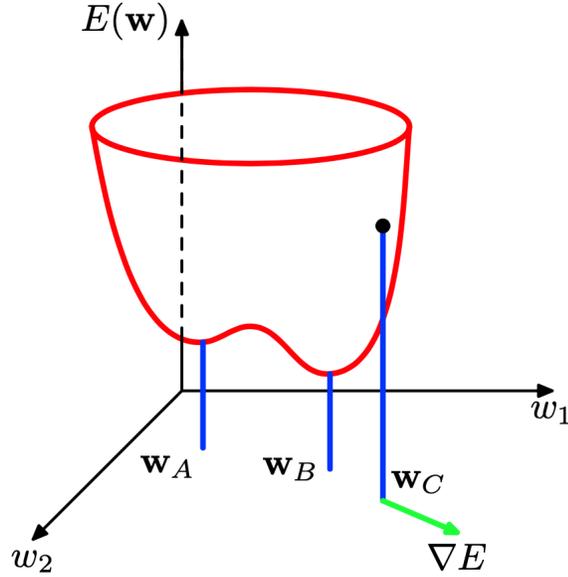


Figura 1.5. Representación de la función de error como una superficie en el espacio de pesos, tomado de [10].

además, cada activación z_j está dada por

$$z_j = h(a_j) \tag{1.11}$$

con $h(\cdot)$ una función de activación. Ahora, tomemos las derivadas de E_n respecto de un peso w_{ji}

$$\frac{\partial E_n}{\partial w_{ji}} = \frac{\partial E_n}{\partial a_j} \frac{\partial a_j}{\partial w_{ji}}. \tag{1.12}$$

Y definimos

$$\delta_j := \frac{\partial E_n}{\partial a_j}. \tag{1.13}$$

Por 1.10 podemos identificar que

$$\frac{\partial a_j}{\partial w_{ji}} = z_i. \tag{1.14}$$

Entonces, sustituyendo δ_j y z_i en 1.12, tenemos que

$$\frac{\partial E_n}{\partial w_{ji}} = \delta_j z_i. \tag{1.15}$$

Haciendo uso de la regla de la cadena nuevamente, tenemos que

$$\delta_j = \frac{\partial E_n}{\partial a_j} = \sum_k \frac{\partial E_n}{\partial a_k} \frac{\partial a_k}{\partial a_j}. \quad (1.16)$$

Esta sumatoria va sobre todas las unidades k a las cuales j se conecta mediante pesos. Finalmente, de la definición de a_j y 1.13, tenemos que

$$\delta_j = h'(a_j) \sum_k w_{kj} \delta_k. \quad (1.17)$$

Es decir, tenemos acceso a los parámetros de las capas anteriores en función de los parámetros de capas posteriores, es por eso que se le llama back propagation, podemos ver una representación esquemática en la figura 1.6.

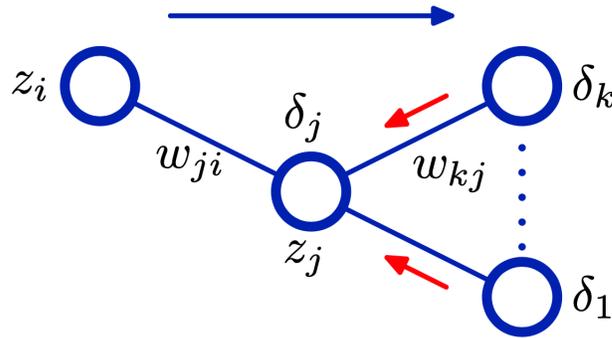


Figura 1.6. Representación del procedimiento de back propagation en una red neuronal. Tomado de [10].

2. EL OBSERVATORIO HAWC

2.1. Características generales

HAWC es un observatorio de rayos gamma ubicado en las faldas del Volcán Sierra Negra en el estado de Puebla, México, a 4100m sobre el nivel del mar. Este consta de 300 detectores Cherenkov en agua distribuidos en un área de 22000 m³ utilizados para detectar de forma indirecta rayos gamma y rayos cósmicos con energías superiores a los 300 GeV.



Figura 2.1. Fotografía del observatorio HAWC. Recuperado de: <https://www.hawc-observatory.org/>.

2.2. Instrumentación y técnica

2.2.1. Efecto Cherenkov

El efecto Cherenkov es un fenómeno donde un medio emite luz como respuesta respuesta al paso de partículas cargadas que se mueven a velocidades más altas que la velocidad de la luz [15]. Para un material genérico de índice de refracción n , la velocidad de la luz en ese material está dada por

$$c_n = c/n, \quad (2.1)$$

con c la velocidad de la luz en el vacío. El medio emite luz en frentes de onda esféricos, los cuales se superponen y dan como resultado una onda cónica, cuyo frente se propaga con un ángulo θ tal que

$$\cos \theta = 1/n. \quad (2.2)$$

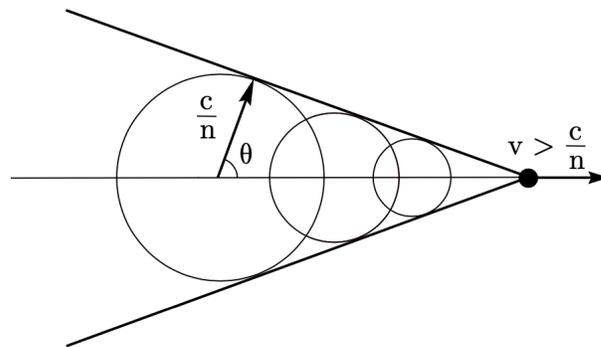


Figura 2.2. Frente de onda de la radiación cherenkov en un medio de índice de refracción n . Tomado de [16].

2.2.2. Tubos fotomultiplicadores (PMTs)

Los tubos fotomultiplicadores son dispositivos electrónicos capaces de detectar fotones, estos funcionan por efecto fotoeléctrico. Estos dispositivos recolectan fotones en el fotocátodo, esto provoca el desprendimiento de electrones por efecto fotoeléctrico que son dirigidos hacia el primer dínodo. Al llegar al dínodo, se desprenden más electrones formando una cascada que finalmente llega al ánodo, obteniendo una señal eléctrica (pulsos) como salida del dispositivo. El ancho de los pulsos que emite el PMT es del orden de los nanosegundos. En la Figura 2.3 podemos observar un esquema de un PMT y un PMT usado en HAWC.

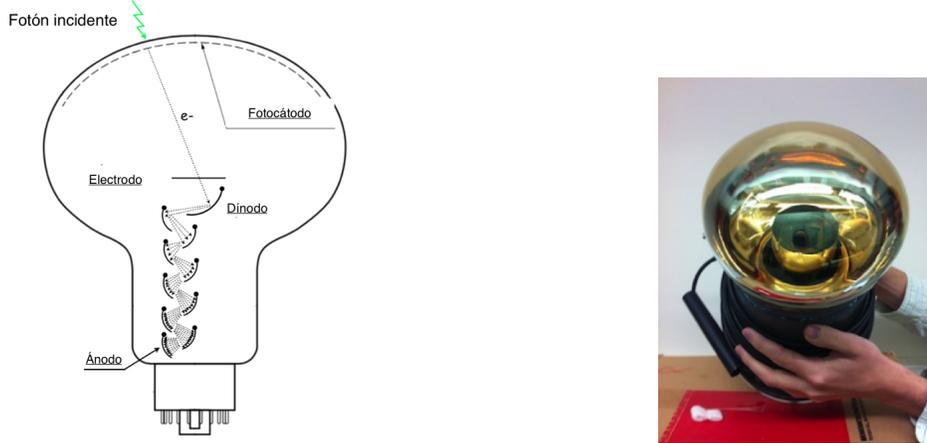


Figura 2.3. A la izquierda: esquema de un PMT. A la derecha: vista física de un PMT utilizado en HAWC. Imágenes tomadas de [16].

2.2.3. Detectores Cherenkov en agua (WCD)

Los detectores de radiación Cherenkov en agua (o WCD por sus siglas en inglés) son tanques cuya función es recolectar la radiación Cherenkov inducida por partículas relativistas con carga eléctrica que los atraviesan. Esta radiación es recolectada por medio de tubos fotomultiplicadores [16]. La información recolectada es digitalizada y se almacena para reconstruir distintas características de los eventos detectados. Los WCD están compuestos por un tanque de acero corrugado de 5 metros de altura y un ancho de 7.3 metros conteniendo aproximadamente 188,000 litros de agua purificada. En el fondo están instrumentados con cuatro PTMs en el fondo con el bulbo apuntando, uno en el centro y tres formando un triángulo. En la figura 2.4 podemos observar una ilustración de un detector Cherenkov usado en HAWC.

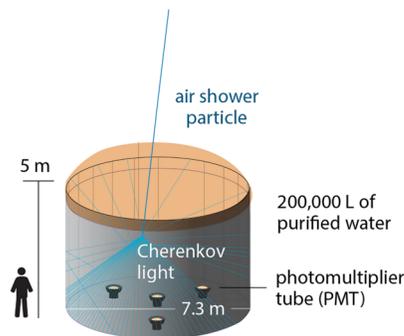


Figura 2.4. Ilustración de un detector Cherenkov usado en HAWC. Imagen tomada del sitio privado de HAWC.

2.3. Software en HAWC

2.3.1. CORSIKA

CORSIKA (COsmic Ray SIMulations for KAscade) es un software dedicado a la simulación de EAS mediante simulaciones Monte Carlo. Este software simula detalladamente el desarrollo de las EAS, permitiendo simular interacciones y decaimientos nucleares, partículas hadrónicas, muones y partículas electromagnéticas en la atmósfera de energías hasta 10^{18} eV [17]. El conjunto de modelos que utiliza CORSIKA es bastante amplio y el uso de estos depende de la energía de los eventos que se requiera simular. Entre los modelos que hay, está Fluka[13] y QGSJET-II [21] para interacciones hadrónicas de baja y alta energía, respectivamente.

2.3.2. HAWCSim

HAWCSim es una aplicación para simular la interacción de las EAS con la materia utilizando [4] para simular la interacción entre las EAS, los detectores y la instrumentación de los mismos. Este software toma las simulaciones Monte Carlo de CORSIKA y simula desde la entrada de las partículas a los WCD, hasta la señal eléctrica emitida por los PMTs.

2.3.3. HAWCSim-Reco

HAWCReco es un programa que simula la recolección de la información proveniente de la instrumentación de los WCD, el sistema de adquisición de datos de HAWC, y reconstruye los parámetros más importantes de las EAS como el núcleo, la energía y dirección. Finalmente, esta información es guardada en archivos de ROOT[11] para su análisis posterior.

3. Metodología

3.1. Entrenamiento de modelos de redes neuronales

Los modelos de redes neuronales fueron definidos y entrenados usando la herramienta TMVA [19] de ROOT [11]. Esta herramienta se especializa en entrenamiento de modelos de machine learning. Permite definir el conjunto de datos a usar en el entrenamiento y la arquitectura del modelo, al finalizar el entrenamiento el modelo óptimo es guardado en un archivo XML. Finalmente, este archivo es utilizado para obtener los pesos calculados durante el entrenamiento de la red neuronal.

3.1.1. Selección de datos de entrenamiento

Para entrenar redes neuronales es necesario contar con un conjunto de datos más relevantes para el problema que deseamos resolver. En nuestro caso, los datos de entrenamiento provienen de simulaciones de EAS. En este trabajo se utilizan los datos reconstruidos por HAWC Sim-Reco.

En la fase de entrenamiento de las redes neuronales se utilizó $2/3$ de los ~ 5 millones de eventos simulados, mientras el tercio restante se utiliza en la fase de pruebas de los modelos. Dichos eventos cumplen las siguientes condiciones:

- Son inicializados por protones, esto se debe a que los protones constituyen un 90 % del espectro de partículas de los rayos cósmicos [14].
- Las partículas inciden isotrópicamente en la Tierra, esto se espera debido a la interacción de los rayos cósmicos con campos magnéticos cuando viajan por el espacio.
- Se simula un flujo con ley de potencia inversa de índice 2.63.

3.1.2. Arquitectura de los modelos entrenados

La arquitectura se define indicando el número de neuronas en cada capa, siendo esta: 14:10:10:1. Es decir, la primera capa está constituida por 14 neuronas, correspondientes a las variables de entrada. En la segunda y tercera capa hay 10 neuronas, encargadas de resolver el problema de reconstruir la energía de los eventos y, finalmente, en la cuarta capa tenemos una neurona de salida, la cual nos indica el resultado de la predicción de la energía del rayo cósmico primario en unidades de $\log(E/GeV)$. A esta estimación le llamamos $\log(E^{reco}/GeV)$ (el superíndice *reco* hace referencia a que es la energía reconstruida por el estimador). La arquitectura se visualiza mejor en la Figura 3.1.

3.1.2.1. Parámetros de entrada para el entrenamiento de modelos

Para entrenar las redes neuronales se utilizaron 14 parámetros de entrada que son los más representativos para la estimación de energía, algunas de estas variables fueron definidas en [20] como variables de entrada para un estimador de energías de rayos gamma. Estos parámetros son construidos como cálculos matemáticos de observables reconstruidos por el software oficial de HAWC (es por esto llevan el prefijo *rec*), estos observables son:

- **rec.nhit**: es el número de PMTs activados durante un evento.
- **rec.ChTot**: es el total de PMTs en operaciones durante ese evento.
- **rec.nTankHit**: es el número de WCDs activados durante un evento. Un WCD se considera activado si al menos un PMT es activado durante el evento.
- **rec.TankTot**: es el total de tanques en operaciones durante ese evento.
- **rec.coreX**: posición en x del núcleo de la cascada.
- **rec.coreY**: posición en y del núcleo de la cascada.
- **rec.LDFamp**: amplitud del ajuste a la LDF de la cascada.
- **rec.zenithAngle**: ángulo cenital de la cascada.
- **rec.fAnnulusCharge0-8**: Porcentaje de carga en el i -ésimo anillo de radio interior $10m*i$ y radio exterior $10m+i$, con i en el intervalo $[0,8]$.

La Tabla 3.1 resume las variables utilizadas para el entrenamiento, la información que contienen sobre las EAS y su definición matemática.

Información contenida sobre las EAS	Descripción	Definición
Multiplicidad	Porcentaje de PMTs activados durante el evento.	$\text{rec.nHit} / \text{rec.nChTot}$
	Porcentaje de tanques Cherenkov activados durante el evento.	$\text{rec.nTankHit} / \text{rec.nTankTot}$
Contenido de carga en el arreglo	Distancia a la que se encontró el núcleo de la cascada, medida desde el centro de HAWC	$\sqrt{(\text{rec.coreX-XHWC})^2 + (\text{rec.coreY-YHWC})^2}$
Distribución lateral y atenuación atmosférica	Amplitud de la distribución lateral del evento.	rec.LDFamp
	Coseno del ángulo cenital de la cascada.	$\cos(\text{rec.zenithAngle})$
	Porcentaje de carga en el i-ésimo anillo de radio interior $10m \cdot i$ y radio exterior $10m \cdot (i+1)$, con i en el intervalo $[0,8]$.	$\text{rec.fAnnulusCharge0-8}$

Tabla 3.1. Parámetros de entrada utilizados y las información que contienen sobre las cascadas atmosféricas extensas. Fuente: elaboración propia.

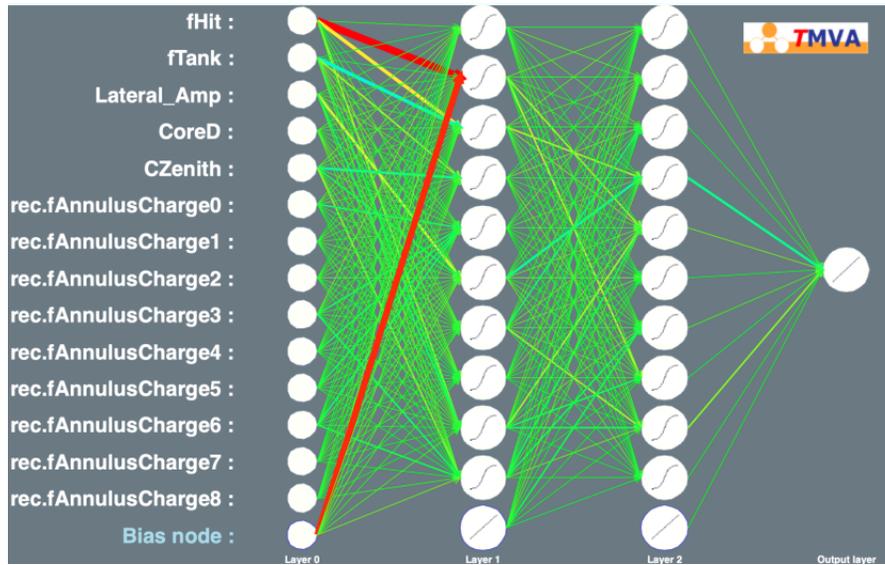


Figura 3.1. Arquitectura de las redes neuronales entrenadas. La arquitectura se define con el número de neuronas en cada capa. En este caso, observamos la arquitectura 14:10:10:1, donde cada círculo blanco representa una neurona y estas están unidas por medio de flechas que representan pesos. Fuente: elaboración propia con TMVA.

3.1.3. Modelos de redes neuronales entrenados

En este proyecto se entrenaron tres modelos de redes neuronales, los cuales comparten la misma arquitectura, pero se diferencian en la cantidad de redes utilizadas en cada caso y en la segmentación de los datos de los eventos reconstruidos para su entrenamiento. La segmentación de los datos se llevó a cabo utilizando dos parámetros clave:

- **coreFiduScale:** Este parámetro indica si el núcleo de la cascada cayó dentro o fuera del arreglo de detectores de HAWC.

- **fhit**: Este parámetro representa el porcentaje de PMTs activados durante un evento.

El primer modelo consiste en una única red neuronal encargada de la reconstrucción de la energía de las partículas primarias, sin distinguir entre eventos cuyo núcleo haya sido localizado dentro o fuera del arreglo de HAWC. En contraste, el segundo modelo está compuesto por dos redes neuronales: una entrenada para reconstruir eventos con el núcleo dentro del arreglo y la otra para eventos con el núcleo fuera de este. Finalmente, el tercer modelo emplea tres redes neuronales, cada una diseñada para reconstruir eventos según el rango de **fhit**: una red para eventos con **fhit** bajo (2.7% - 22%), otra para eventos con **fhit** medio (22% - 47%), y una tercera red para eventos con **fhit** \geq 47%. Las características de estos modelos se detallan en la Tabla 3.2.

Modelo	Número de redes neuronales usadas	fhit bin	Dentro/fuera de HAWC
Primer modelo	1	[2.7%, 100 %]	Ambos
Segundo modelo	2	[2.7%, 100 %]	Fuera
		[2.7%, 100 %]	Dentro
Tercer modelo	3	(2.7% - 22 %]	Ambos
		(22% -47 %]	
		(47%-100 %]	

Tabla 3.2. Resumen de los modelos entrenados, el número de redes neuronales usadas, el detalle de los bins de **fhit** de cada red neuronal y si se usaron datos de núcleos encontrados dentro o fuera de HAWC.

3.2. Cortes de calidad y cortes de selección de eventos

Los cortes de calidad y cortes de selección de eventos son filtros aplicados tanto a eventos reconstruidos como a simulaciones reconstruidas, estos cortes fueron definidos en [6] y son un estándar para el análisis de reconstrucción de energías.

Los cortes de calidad discriminan los eventos con las siguientes características:

- el núcleo de la cascada no fue correctamente reconstruido y
- el ángulo de la cascada no fue correctamente reconstruido.

Por otro lado, los cortes de selección se encargan de aislar eventos para reducir errores sistemáticos propios de la reconstrucción de eventos hadrónicos. Estos cortes son:

- Los eventos deben pasar un umbral de multiplicidad, activando un mínimo de 75 PMTs dentro del arreglo.
- Se escogen eventos con ángulo cenital $\theta \leq 16.7$, correspondiendo a cascadas verticales.
- Al menos 40 PMTs deben ser activados en un radio de 40 metros alrededor del núcleo de la cascada. Este criterio asegura que el núcleo haya caído dentro del arreglo o a máximo 15 metros del mismo.

4. RESULTADOS

4.1. Comparación entre modelos

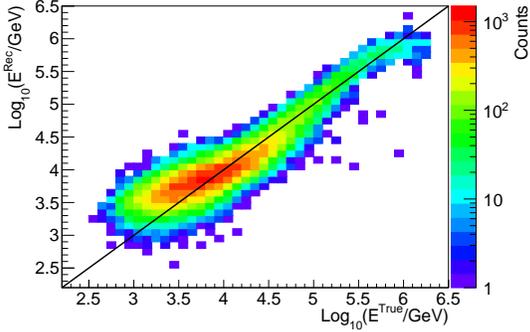
En esta sección se comparan los tres modelos de redes neuronales. Durante la fase de entrenamiento, se utilizó dos tercios de los datos para entrenar los modelos, reservando el tercio restante para la fase de pruebas. En la fase de pruebas, se reconstruyó la energía con los modelos entrenados y se comparó con la energía real, utilizando $\log(E_{True}/GeV)$. Esta comparación se hizo mediante histogramas 2D que muestran la energía reconstruida en función de la energía verdadera y el conteo de eventos en una escala de colores, a los cuales se aplicaron cortes de calidad para evitar errores sistemáticos. El resultado ideal es que los eventos se distribuyan en la diagonal identidad, lo que indicaría que la reconstrucción fue exitosa. Sin embargo, el comportamiento esperado es que los eventos se distribuyan alrededor de esta línea con cierta dispersión.

4.1.1. Comparación entre modelos de redes neuronales

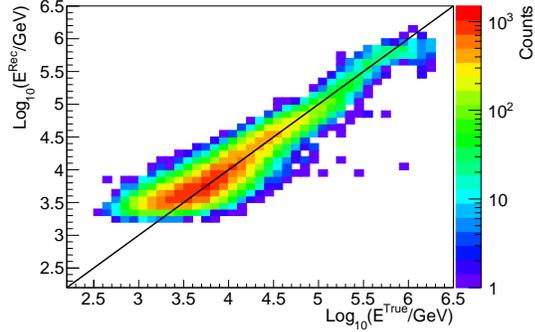
En la Figura 4.1 se presenta este análisis para los tres modelos de redes neuronales entrenados. Aquí observamos que los tres modelos presentan mayor dispersión en la región de bajas energías ($\text{Log}(E/GeV) < 4$). Este comportamiento mejora en la región de medianas energías ($4 \leq \text{Log}(E/GeV) < 5.5$), donde podemos observar una mejor distribución de eventos alrededor de la identidad, presentando una mayor linealidad en los tres estimadores. Finalmente, en la región de altas energías ($\text{Log}(E/GeV) > 5.5$) los primeros dos modelos tienden a subestimar la energía, mientras que el tercer modelo muestra una mejor linealidad.

En resumen, los tres modelos entrenados muestran consistencia al reconstruir bajas y medianas energías, aumentando su linealidad en estos rangos. Sin embargo, en las altas energías, los primeros dos modelos pierden linealidad, mientras que el tercer modelo la mantiene, demostrando ser el más adecuado para reconstruir la

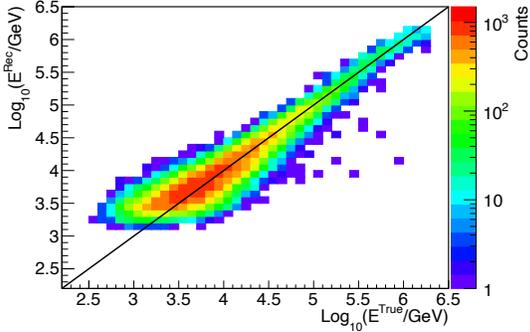
energía.



(a) Primer modelo



(b) Segundo modelo



(c) Tercer modelo

Figura 4.1. Se muestra un histograma 2D de la energía reconstruida vs la energía verdadera de los tres modelos entrenados en este trabajo (a,b y c, respectivamente). Se muestra la identidad como una línea negra en la diagonal. Fuente: elaboración propia.

4.1.2. Comparación del tercer modelo contra el estimador actual de HAWC

En la sección 4.1.1 se comparó el desempeño de los tres modelos de redes neuronales entrenados en este trabajo, de los cuales el tercero fue el que presenta una mejor reconstrucción de la energía, este modelo será llamado "la red neuronal" de aquí en adelante. En esta sección se comparará la red neuronal contra el estimador oficial de HAWC.

Para cuantificar y comparar las calidad de las reconstrucciones de energía definimos el bias como

$$\text{bias} = \langle \Delta \log(E) \rangle = \langle \log_{10}(E_{\text{Rec}}/\text{GeV}) - \log_{10}(E_{\text{True}}/\text{GeV}) \rangle. \quad (4.1)$$

En este trabajo se calculó el bias de la energía reconstruida por cada cuarto de década de la energía verdadera. Podemos visualizar un ejemplo del calculo del bias en la Figura 4.2, donde este se obtiene por medio de un ajuste gaussiano al histograma de $\Delta \text{Log}(E)$ en el intervalo de $\text{Log}(E^{\text{true}}/\text{GeV})$ entre 3.50 – 3.75, donde el bias es 0.035.

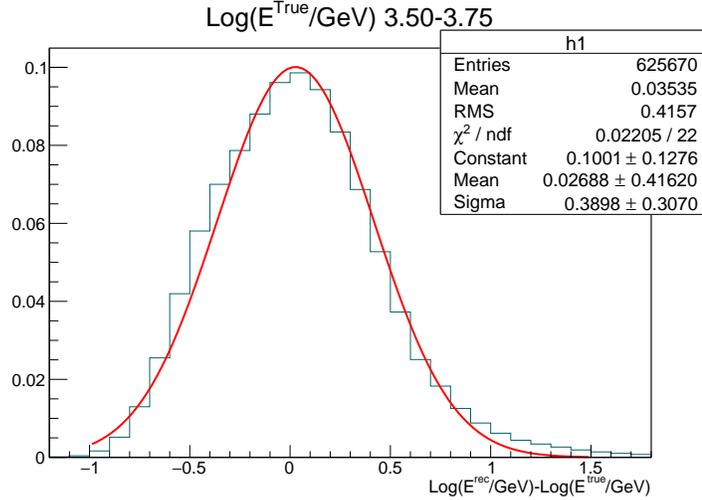


Figura 4.2. Histograma normalizado de $\Delta \text{Log}(E)$ en el intervalo (3.50-3.75) de $\log(E^{\text{True}}/\text{GeV})$ y un ajuste gaussiano en rojo.

Las Figuras 4.3 y 4.4 ilustran el comportamiento de la energía reconstruida y el bias, respectivamente. En la región de bajas energías ambos estimadores presentan dificultades en la reconstrucción, esto lo podemos notar en los eventos dispersos alrededor de la diagonal en la Figura 4.3, la Figura 4.4 nos confirma este comportamiento, puesto que es donde el bias tiene valores más alejados de cero. En el rango de energías medias, ambos estimadores logran una excelente reconstrucción de la energía. Sin embargo, en las altas energías, los estimadores empiezan a divergir: el bias del estimador de HAWC se aleja rápidamente de cero, mientras que el de la red neuronal se mantiene cercano a este valor, indicando una mejor reconstrucción. Este comportamiento es nuevamente coherente con los histogramas 2D mostrados, donde en las altas energías vemos una mayor linealidad en la red neuronal, mientras el estimador de HAWC subestima la energía.

En resumen, los estimadores tienen un comportamiento similar en los intervalos de bajas y medias energías. Sin embargo, en altas energías la red neuronal muestra una mayor linealidad según la distribución del bias en 4.4 y la distribución de eventos alrededor de la diagonal identidad en los histogramas 2D en 4.3.

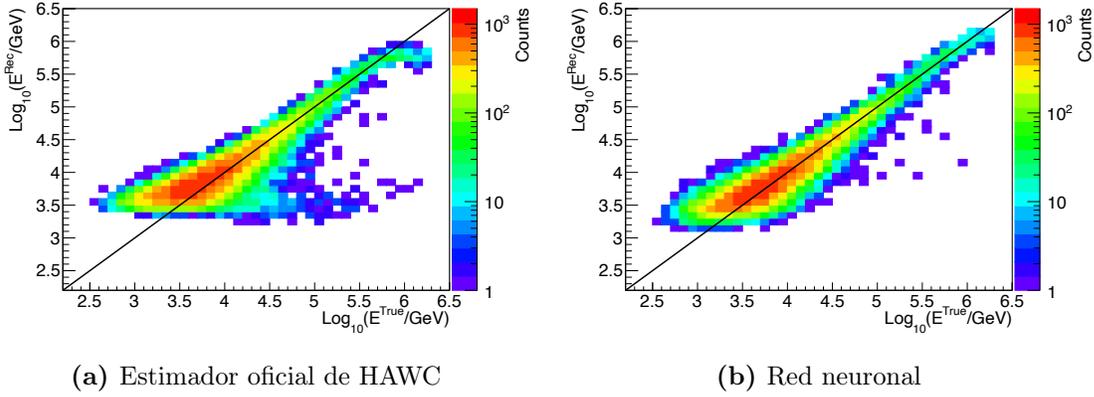


Figura 4.3. Histogramas 2D de la energía reconstruida por el estimador oficial de HAWC (a) y por el mejor modelo de redes neuronales (b). Fuente: elaboración propia.

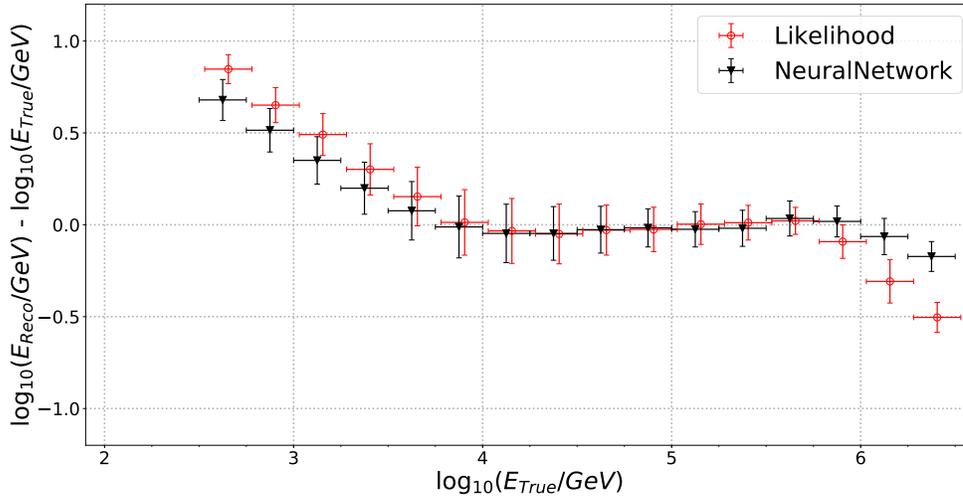


Figura 4.4. Distribución del bias del estimador actual de HAWC (Likelihood) y el tercer modelo entrenado en este trabajo (NeuralNetwork). Fuente: elaboración propia.

4.2. Pruebas en eventos iniciados por núcleos de hierro

Para probar la robustez de la red neuronal, se reconstruyó la energía de eventos iniciados por núcleos atómicos de hierro. Esto nos permite evaluar la reconstrucción en eventos diferentes a los utilizados para entrenar nuestros modelos. La gráfica 4.6 nos muestra que el tercer modelo no es tan robusto a energías medias, sin embargo, reconstruye de buena manera en altas energías, siendo consistente con las reconstrucciones de la energía de EAS iniciadas por protones. La distribución del bias es consistente con lo que muestra la figura 4.5, donde observamos puntos por debajo

de la línea identidad en gran parte de la figura 4.5a. Por otro lado, la reconstrucción del tercer modelo mostrado en 4.5b también presenta malas reconstrucciones a energías medias, sin embargo, estas se van corrigiendo a medida que crece la energía, nuevamente siendo consistente con 4.6.

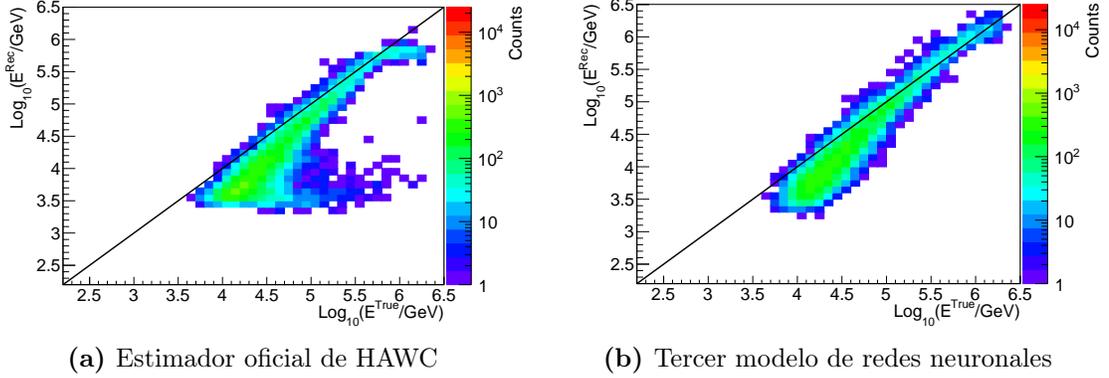


Figura 4.5. Se muestra un histograma 2D ilustrando la energía verdadera contra la energía reconstruida por el estimador oficial de HAWC (a), y la energía reconstruida por el tercer modelo de redes neuronales (b) utilizando EAS inducidas por núcleos de hierro. Fuente: elaboración propia.

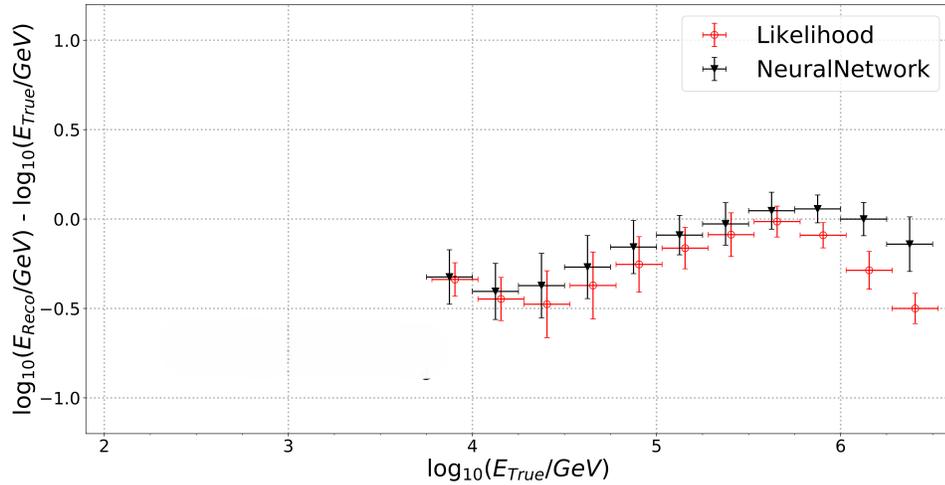


Figura 4.6. Comparación del bias de la energía reconstruida por el estimador actual de HAWC y el tercer modelo de redes neuronales para eventos iniciados por núcleos de hierro. Fuente: Elaboración propia.

4.3. Discusión y conclusiones

En este trabajo, se entrenaron, evaluaron y compararon tres modelos diferentes de redes neuronales utilizando simulaciones Monte Carlo de cascadas atmosféricas

iniciadas por protones. Cada modelo se distinguió por la segmentación de los datos empleados en el entrenamiento. El análisis mostró que el modelo más efectivo fue aquel que segmenta los datos en función del parámetro `fhit`, dividiendo los eventos en bins de bajo, medio y alto `fhit`. Este modelo superó a los otros dos, que se basaban en discriminar eventos según la posición del núcleo dentro o fuera del arreglo de WCDs. Al comparar el desempeño de este tercer modelo con el estimador oficial de HAWC, se encontró que, mientras ambos estimadores ofrecen resultados similares en bajas y medianas energías, el modelo de redes neuronales muestra un sesgo significativamente menor en energías altas, lo que sugiere una mejor reconstrucción en ese rango de energías.

Además de evaluar el rendimiento con protones, se exploró la robustez del modelo aplicándolo a la reconstrucción de eventos iniciados por núcleos de hierro. Aunque el modelo mantuvo una buena precisión en altas energías, en energías medias el sesgo fue considerablemente mayor. Este resultado resalta la importancia de diversificar los datos de entrenamiento, sugiriendo que el uso de otras especies de rayos cósmicos podría mejorar la generalización del modelo. Asimismo, podría ser beneficioso considerar diferentes estrategias de segmentación, como un esquema basado en los ángulos de arribo de las cascadas al observatorio, para mejorar aún más la reconstrucción de energía.

En resumen, los resultados de esta tesis confirman que un modelo de red neuronal entrenado con una segmentación adecuada de los datos ofrece mejoras sustanciales en la reconstrucción de energía, particularmente en eventos de alta energía. Sin embargo, también se identificaron áreas donde el modelo podría beneficiarse de ajustes adicionales en el proceso de entrenamiento y segmentación de datos, lo que abre la puerta a futuras investigaciones y desarrollos en esta área.

BIBLIOGRAFÍA

- [1] ABEYSEKARA, A. U. y otros: «All-sky Measurement of the Anisotropy of Cosmic Rays at 10 TeV and Mapping of the Local Interstellar Magnetic Field». , 2019, **871(1)**, 96. doi: 10.3847/1538-4357/aaf5cc.
- [2] —: «The High-Altitude Water Cherenkov (HAWC) observatory in México: The primary detector». *Nuclear Instruments and Methods in Physics Research A*, 2023, **1052**, 168253. doi: 10.1016/j.nima.2023.168253.
- [3] ABU-MOSTAFA, YASER S.; MAGDON-ISMAIL, MALIK y LIN, HSUAN-TIEN: *Learning From Data*. AMLBook, 2012.
- [4] AGOSTINELLI, S. y otros: «GEANT4—a simulation toolkit». *Nucl. Instrum. Meth. A*, 2003, **506**, pp. 250–303. doi: 10.1016/S0168-9002(03)01368-8.
- [5] ALBERT, A. y otros: «Cosmic ray spectrum of protons plus helium nuclei between 6 and 158 TeV from HAWC data». *Phys. Rev. D*, 2022, **105**, p. 063021. doi: 10.1103/PhysRevD.105.063021.
<https://link.aps.org/doi/10.1103/PhysRevD.105.063021>
- [6] ALFARO, R. y otros: «All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV». , 2017, **96(12)**, 122001. doi: 10.1103/PhysRevD.96.122001.
- [7] ALVES BATISTA, RAFAEL; BITEAU, JONATHAN; BUSTAMANTE, MAURICIO; DOLAG, KLAUS; ENGEL, RALPH; FANG, KE; KAMPERT, KARL-HEINZ; KOSTUNIN, DMITRIY; MOSTAFA, MIGUEL; MURASE, KOHTA; OIKONOMOU, FOOTEINI; OLINTO, ANGELA V.; PANASYUK, MIKHAIL I.; SIGL, GUENTER; TAYLOR, ANDREW M. y UNGER, MICHAEL: «Open Questions in Cosmic-Ray Research at Ultrahigh Energies». *Frontiers in Astronomy and Space Sciences*, 2019, **6**. ISSN 2296-987X. doi: 10.3389/fspas.2019.00023.
<http://dx.doi.org/10.3389/fspas.2019.00023>

- [8] ARTEAGA-VELÁZQUEZ, J. C. y otros: «Estimate of the energy spectrum of the light component of cosmic rays in HAWC using the shower age and the fraction of hit PMT's». En: *35th International Cosmic Ray Conference (ICRC2017)*, volumen 301 de *International Cosmic Ray Conference*, p. 501, 2017. doi: 10.22323/1.301.0501.
- [9] BEATTY, JAMES J. y WESTERHOFF, STEFAN: «The Highest-Energy Cosmic Rays». *Annual Review of Nuclear and Particle Science*, 2009, **59(1)**, pp. 319–345. doi: 10.1146/annurev.nucl.58.110707.171154.
<https://doi.org/10.1146/annurev.nucl.58.110707.171154>
- [10] BISHOP, CHRISTOPHER M.: *Pattern Recognition and Machine Learning*. Springer, 2006.
- [11] BRUN, RENE y RADEMAKERS, FONS: «ROOT — An object oriented data analysis framework». *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1997, **389(1)**, pp. 81–86. ISSN 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
<https://www.sciencedirect.com/science/article/pii/S016890029700048X>
- [12] DE ANGELIS, A. y PIMENTA, M.: *Introduction to Particle and Astroparticle Physics: Multimessenger Astronomy and its Particle Physics Foundations*. Undergraduate Lecture Notes in Physics. Springer International Publishing, 2018. ISBN 9783319781808.
<https://books.google.com.gt/books?id=JaQHtAEACAAJ>
- [13] FERRARI, A y otros: «FLUKA: A multi-particle transport code (program version 2005)». *Informe técnico CERN-2005-010, SLAC-R-773, INFN-TC-05-11, CERN-2005-10*, Geneva, 2005. doi: 10.2172/877507.
<https://www.osti.gov/biblio/877507>
- [14] GAISSER, T.K.: *Cosmic Rays and Particle Physics*. Cambridge University Press, 1991. ISBN 9780521339315.
<https://books.google.com.gt/books?id=5JMmJI7Mm30C>
- [15] GRIEDER, PETER: *Extensive air showers. High energy phenomena and astrophysical aspects. A tutorial, reference manual and data book. Vol. I and II*. volumen 1, 2010. ISBN 978-3-540-76940-8. doi: 10.1007/978-3-540-76941-5.

- [16] HAMPEL-ARIAS, ZIGFRIED: *Cosmic Ray Observations at the TeV Scale with the HAWC Observatory*. Tesis doctoral, University of Wisconsin-Madison, 2017.
- [17] HECK, D.; y otros: «CORSIKA: a Monte Carlo code to simulate extensive air showers.» *Informe técnico Report No. FZKA-6019*, 1998.
- [18] HEITLER, W.: *The Quantum Theory of Radiation*. Dover, New York, third^a edición, 1984.
- [19] HOECKER, ANDREAS y otros: «TMVA: Toolkit for Multivariate Data Analysis». , 2007, (**ACAT**), p. 40.
- [20] MARINELLI, SAMUEL STEPHENS: *CONSTRAINTS ON LORENTZ-INVARIANCE VIOLATION WITH THE HAWC OBSERVATORY*. Phd thesis, Michigan State University, 2019.
- [21] OSTAPCHENKO, S.: «Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model». , 2011, **83(1)**, 014018. doi: 2017ICRC_CREEP410.1103/PhysRevD.83.014018.