

Title of Document

Name of Author

November 23, 2023

Metodología

El aprendizaje por transferencia es una técnica fundamental en el campo de la inteligencia artificial que permite aprovechar el conocimiento adquirido en una tarea previa para mejorar el rendimiento en una tarea relacionada o diferente. En lugar de entrenar un modelo desde cero para cada nueva tarea, el aprendizaje por transferencia permite reutilizar modelos, llamados "modelos base" o "preentrenados", que han sido entrenados en grandes conjuntos de datos y han aprendido características generales de los datos. Estos modelos preentrenados se ajustan luego a la tarea específica mediante un proceso de ajuste fino o *fine-tuning*, lo que ahorra tiempo y recursos de entrenamiento.

La elección de este paradigma de inteligencia artificial es fundamental ya que los isótopos calculados experimentalmente provenientes de la *Atomic Mass Evaluation 2020* (Huang et al., 2021) suman un total de 2525, lo que dificulta que un paradigma tradicional de machine learning prediga los datos aproximados a los experimentales.

En este sentido, el presente trabajo propone una alternativa basada en *transfer learning*, la cual consiste en entrenar una red neuronal con los datos obtenidos mediante el modelo de la gota líquida y tomar esta red neuronal como modelo preentrenado para la generación de un nuevo entrenamiento con los datos experimentales, con el objetivo de llegar a la predicción de la energía de enlace del nucleón.

Tratamiento de datos

Antes de iniciar el proceso de entrenamiento, se lleva a cabo el preprocesamiento de los datos, que implica la consolidación de conjuntos de datos. En este procedimiento, se fusionan los datos teóricos derivados del proceso de la gota líquida, como se describe en el trabajo de (Pastore & Carmini, 2021), con los datos experimentales obtenidos de la *Atomic Mass Evaluation 2020* (Huang et al., 2021). La unificación de estos conjuntos de datos se realiza mediante los valores de Symbol, Z (*número atómico*), y N (*número másico*), con el objetivo de construir un

conjunto de datos más completo y robusto. En este contexto, se excluyen los datos correspondientes a los elementos Hidrógeno, Helio y Litio, dados sus pesos atómicos extremadamente bajos, según (Gao et al., 2021), lo que los excluye del modelo considerado. Este proceso garantiza la coherencia y relevancia de los datos utilizados en el entrenamiento del modelo, al tiempo que se ajusta a las consideraciones teóricas y experimentales previamente establecidas.

Se procede con la especificación de las variables que serán empleadas en el cálculo de los datos de energía de enlace, siguiendo el modelo de la Gota Líquida. Estos valores se derivan de la revisión literaria, donde se han identificado los siguientes parámetros: constante volumétrica (15.782), constante superficial (17.9042), constante de Coulomb (0.72404), constante de asimetría (23.7193) y constante de paridad (12). Estas magnitudes desempeñan un papel fundamental en la formulación del modelo, y su incorporación permite establecer las relaciones pertinentes en el cálculo de la energía de enlace, asegurando así la coherencia y precisión del modelo según los principios teóricos y datos experimentales disponibles en la literatura especializada.

Utilizando el proceso de obtención de las dimensiones del conjunto de datos, se emplea un método iterativo para calcular las variables subsiguientes, las cuales están detalladas en la Tabla 1. El objetivo subyacente es la obtención de un conjunto de datos más robusto y con un mayor número de variables, optimizando así su complejidad y enriqueciendo las características disponibles para el entrenamiento preliminar del modelo. Este enfoque iterativo permite una expansión controlada y sistemática del conjunto de datos, asegurando que las nuevas variables incorporadas contribuyan de manera significativa a la representación del fenómeno subyacente, promoviendo así un entrenamiento más efectivo y comprensivo del modelo.

Nomenclatura	Detalle	Obtención
Symbol	Símbolo del elemento	Obtenido de tabla periódica
Z	Número de protones	Obtenido de tabla periódica
N	Número de neutrones	Obtenido de tabla periódica
A	Número Másico	$A = Z + N$
χ	Cociente entre N y P	$\chi = \frac{N}{Z}$
M	Número mágico	$M(Z, N) = \begin{cases} 1, & \text{si } Z \parallel N \in \{2, 8, 20, 28, 50, 82, 126\}, \\ 0, & \text{caso contrario} \end{cases}$
δ	Paridad	$\delta = \frac{(-1)^Z + (-1)^N}{2}$
BEA	Energía de enlace por nucleón	$BEA = a_v - a_s/A^{1/3} - \frac{a_c(Z^2 - Z)}{A^{4/3}} - \frac{a_a(N - Z)^2}{A^2} + \frac{a_p}{A^{3/2}}\delta$

Table 1: Detalle de datos

El conjunto de datos, tras el preprocesamiento de información, se almacena con un total de 7310 filas. Posteriormente, se lleva a cabo un reordenamiento aleatorio de la información. Inicialmente, se asigna el 10% de los datos, equivalente a 731 filas, para constituir el conjunto de pruebas, dejando un remanente de 6579 filas. De este remanente, se destina el 80% para configurar el conjunto de entrenamiento, totalizando 5263 filas, mientras que el restante 20% conforma el conjunto de validación, con un total de 1316 filas. Ambos conjuntos, de en-

trenamiento y validación, son almacenados en sus respectivos directorios para su posterior uso en las fases correspondientes del proceso. Este procedimiento asegura la disponibilidad de datos bien estructurados y distribuidos estratégicamente para la fase de entrenamiento y evaluación del modelo.

Arquitectura y Entrenamiento

Después de la adquisición y clasificación de los datos, se inicia el proceso de selección de la arquitectura de Redes Neuronales Artificiales (ANN) que será empleada en el modelado. Previamente a la determinación de la ANN final, se llevan a cabo diversas pruebas exploratorias de arquitecturas con el propósito de evaluar su desempeño y determinar cuál de ellas se adapta de manera óptima a la resolución del problema planteado. Estas pruebas implican la variación sistemática de la topología y configuración de capas de la ANN, siendo los resultados detallados en la tabla de experimentos 2. La elección de la arquitectura óptima se basa en métricas específicas relacionadas con el rendimiento y la generalización del modelo, siendo esenciales para determinar la eficacia y adecuación de cada configuración probada. Este enfoque riguroso asegura la selección de la ANN que mejor se ajusta a las complejidades inherentes al problema y optimiza su capacidad de aprendizaje y generalización.

Hyperparameters	Activation Funtion	Input	Layers	Output Funtion	Output Data
$epoch = 1500$	RELU	Z	$layer = 3$	RELU	$BEA\ MSE = 0.0304$
$batch_size = 32$		N			
$optimizer = ADAM$		A			
$lr = 0.001$		χ			
$kernel_regularizer = L_2$		M			
$L_2 = 0.01$		δ			

Table 2: Entrenamiento