

# ENERGIAS DE ENLACE NUCLEAR EN REDES NEURONALES ARTIFICIALES

Autor: Esteban Kuzmicich

Fuente: Lin-Xing Zeng, Yu-Ying Yin, Xiao-Xu Dong, and Li-Sheng Geng. Phys. Rev. C 109, 034318 (2024)



# MOTIVACION

Las energías de enlace nuclear son **inputs esenciales** para diversas áreas de estudio de la física nuclear, como la síntesis de núcleos pesados y estudios de astrofísica nuclear.

Ejemplos:

## Síntesis de núcleos ultra pesados.

Se hacen chocar núcleos para obtener un núcleo con un  $Z$  mayor

## r-process:

Captura rápida de neutrones

Supernovas y estrellas de neutrones

Alta densidad neutrónica y temperatura

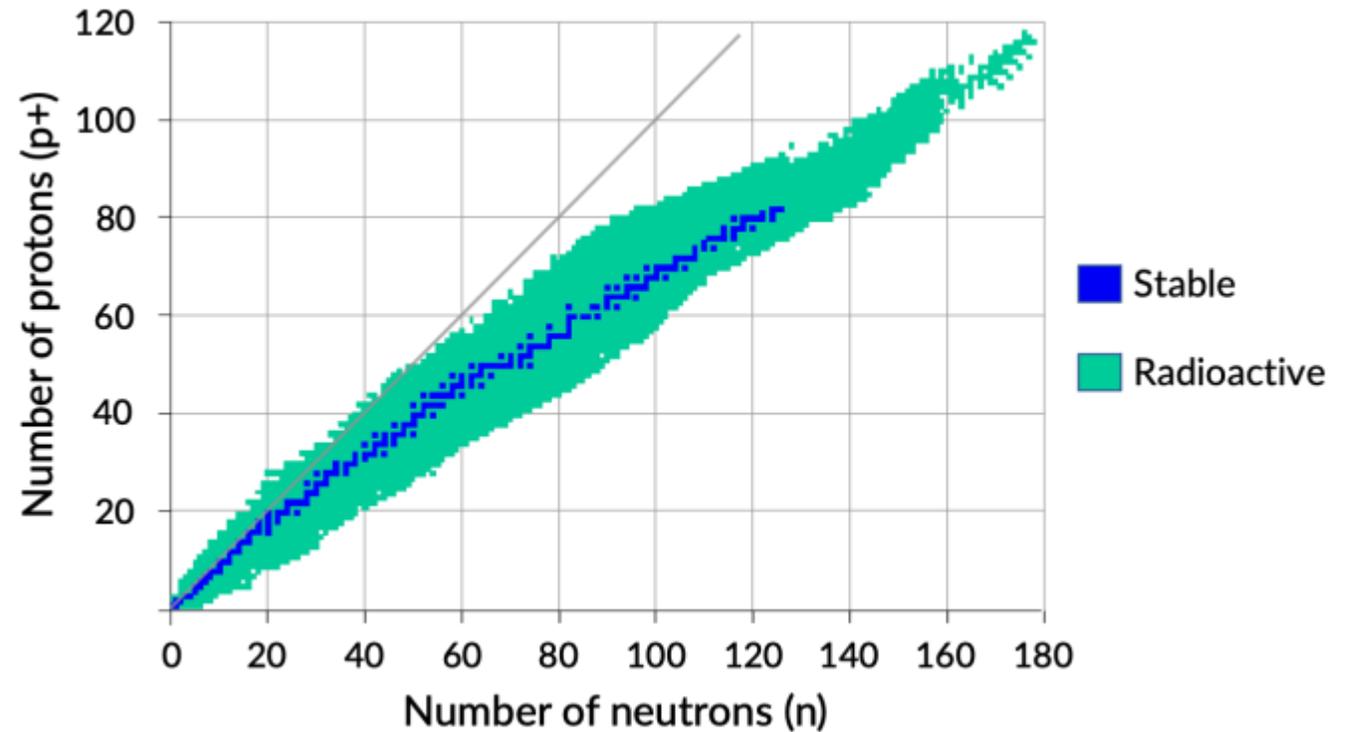
Formación de núcleos pesados

## X-ray bursts:

Estrella de neutrones acumula gas (He y H) de una estrella cercana

Se genera una gran fusión nuclear por alta  $P$  y  $T$

Se liberan Rayos X



## Introducción

-En el último AME2020 se compilaron las masas de 3556 núcleos.

-Modelos utilizados para el cálculo de las BEs:

the relativistic mean field model (RMF)

Hartree-Fock-Bogoliubov model (HFB)

the finite-range droplet model (FRDM)

Weizsäcker-Skyrme model (WS)

Etc.

-Todos estos modelos describen las masas del AME con una buena precisión, entre 0.3 MeV y 1-2 MeV.

El más exitoso de ellos: WS4, RMSD de 0.298 MeV.

Siendo el primer término la fórmula ya vista para BE en la materia. El segundo término  $E_{mic}$  es la corrección de Skyrme, corrige por capas, da estabilidad a los números mágicos.

El modelo WS se formula como una **energía total nuclear** compuesta por dos partes:

$$E_{tot}(Z, A, \{\beta\}) = E_{LD}(Z, A, \{\beta\}) + \Delta E_{mic}(Z, A, \{\beta\})$$

$$E_D(A, Z) = a_v A - a_s \frac{Z^2}{A^{1/3}} - \frac{(N-Z)^2}{A} + \delta(A, Z)$$

- ELD: energía macroscópica tipo **gota líquida deformada**.
- $\Delta E_{mic}$ : correcciones **microscópicas** (efectos de capa, emparejamiento, etc.).

# FORMALISMO TEÓRICO

- ¿Qué es una red neuronal?

Una **red neuronal** es un modelo matemático inspirado en cómo funciona el cerebro humano, diseñado para **procesar información, aprender de datos y hacer predicciones o clasificaciones**.

Estructura:

La unidad básica de las ANN son las **neuronas**, cada una de ellas recibe uno o varios números de entrada o **inputs  $I_j$** , lo multiplica por un **peso  $W_j$** , le suma un **sesgo  $b_j$**  y lo pasa por una **función de activación  $f$** , dando como resultado un dado Output.

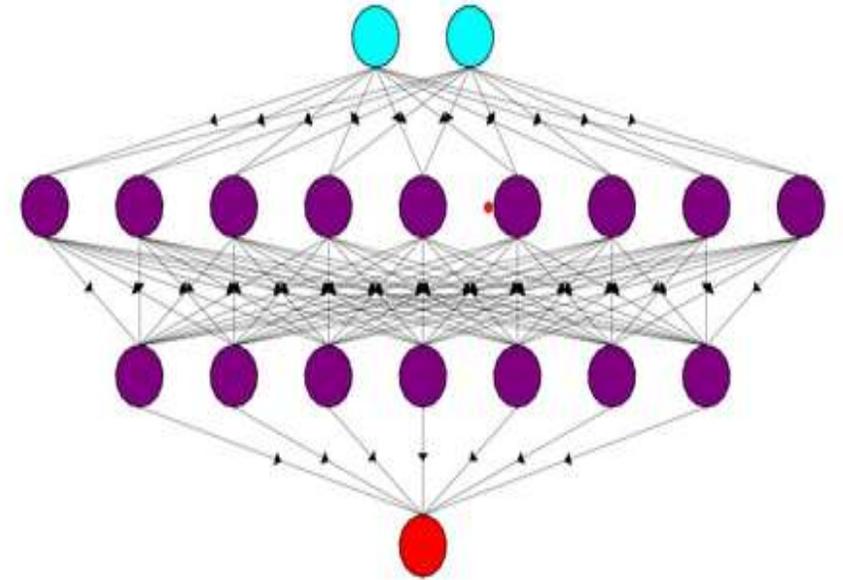
$$O_h = f(W_j \cdot I_j + b_j).$$

Las mismas se organizan en capas:

Capa de entrada, recibe los inputs.

Capa/s oculta/s, procesan la información a través de combinaciones y transformaciones.

Capa de salida: devuelve un resultado, en este caso un valor numérico.



## EJEMPLO DE RED NEURONAL

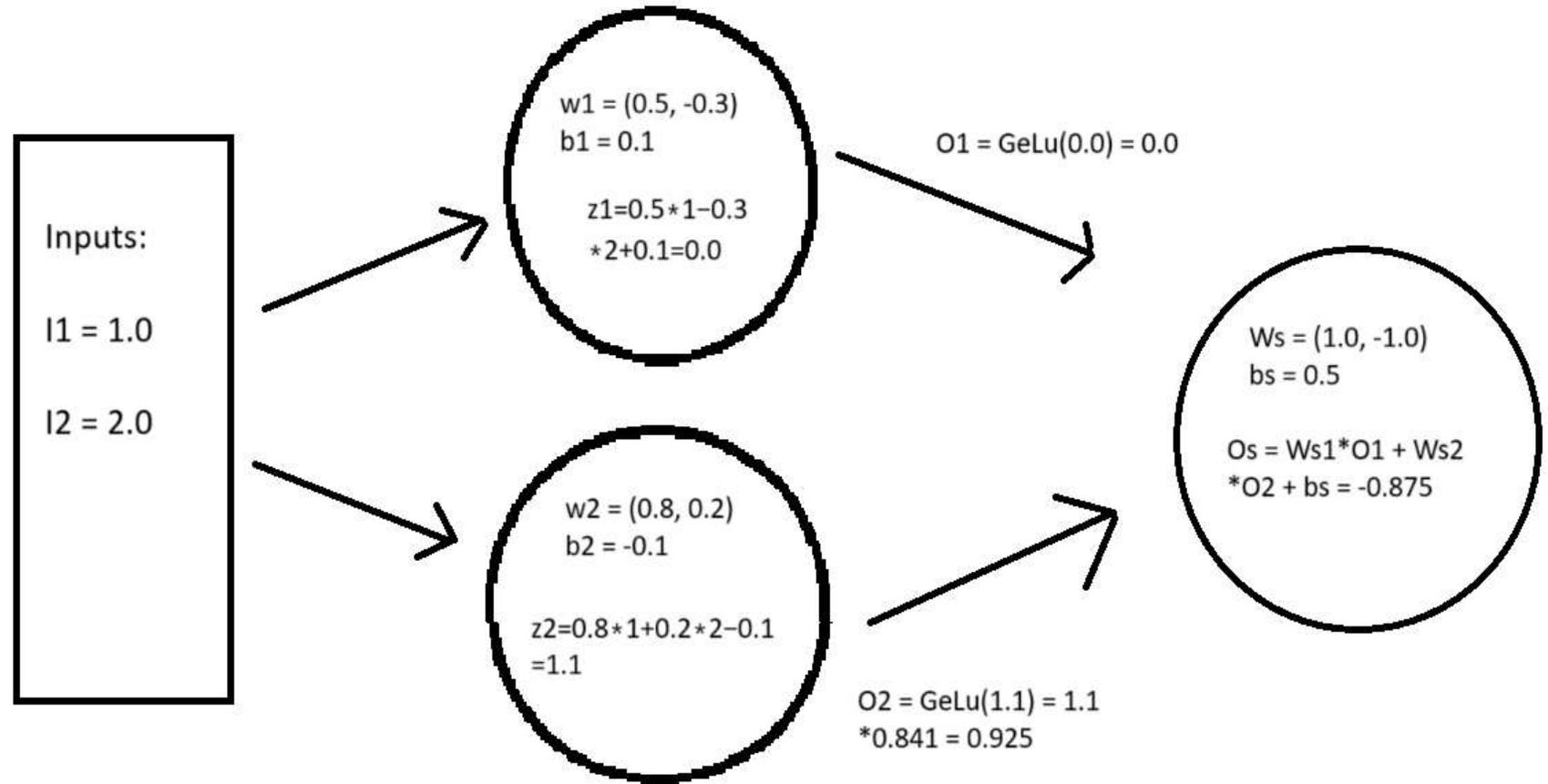
- Red simple de una capa oculta:
- Capa oculta:** 2 neuronas
- Función de activación capa oculta:** GeLU (Gaussian Error Linear Unit)
- Función de activación capa salida:** identidad

$$\mathbf{O} = \text{GeLU}(\mathbf{W} \cdot \mathbf{I} + \mathbf{b}).$$

Función GeLU:

$$\text{GeLU}(x) = \frac{x}{2} \left[ 1 + \text{erf} \left( \frac{x}{\sqrt{2}} \right) \right]$$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$



Se compara el valor resultante de la ANN con el dato, se calcula un error con una función pérdida L.

La ANN **reajusta los pesos** en cada neurona, utilizando dos métodos a continuación.

Back propagation: **calculamos la variación de pérdida respecto a cada peso y sesgo**. Regla de la cadena.  
 $a = f(z)$ .

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w}$$

- $\frac{\partial L}{\partial a}$  viene de la capa siguiente (o de la salida si es la última capa).
- $\frac{\partial a}{\partial z} = f'(z)$  es la derivada de la función de activación.
- $\frac{\partial z}{\partial w} = x$ , la entrada a esa neurona.

Para sesgos  $b$ :

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial a} \cdot f'(z)$$

Gradient descent:

Ajustamos los pesos y sesgos.

$\eta$  es la **tasa de aprendizaje**.

Esto “mueve” los pesos en la **dirección que reduce la pérdida**.

$$w_f = w_i - \eta \frac{\partial L}{\partial w_i}$$

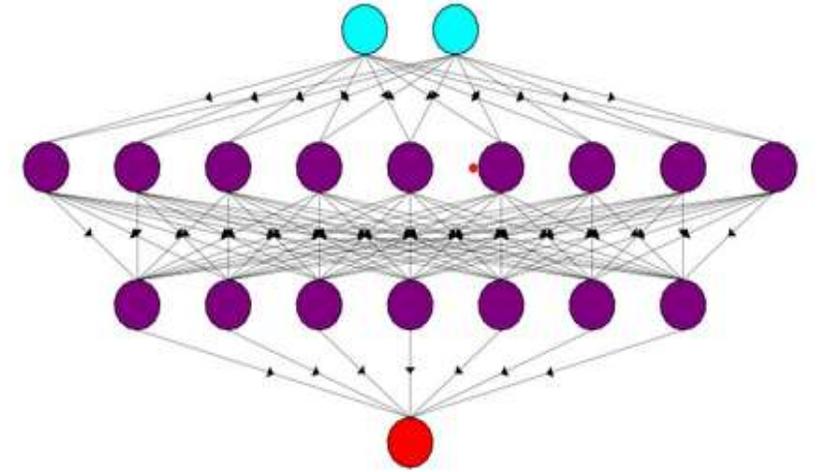
$$b_f = b_i - \eta \frac{\partial L}{\partial b_i}$$

## RED NEURONAL UTILIZADA

En el paper se utilizaron 4 redes neuronales distintas, ANN2, ANN4, ANN6 y ANN7.

Características:

- 1) Una capa de entrada con 2, 4, 6 o 7 inputs.
- 2) Dos capas ocultas con 9 y 7 neuronas respectivamente. Función de activación: GeLU
- 3) Una capa de output con una sola neurona. Función de activación: identidad
- 4) Tasa de aprendizaje  $\eta = 0.0001$
- 5) Función pérdida utilizada: LOSS
- 6) Se utilizó el Adam algorithm, una variante del metodo gradient descent.
- 7) Inputs: Z, N, Zeo, Neo,  $\Delta Z$ ,  $\Delta N$  y ASY
- 8) Los vectores de pesos y sesgos fueron inicializados con las mismas semillas random para todas las ANN.



$$\text{LOSS} = \frac{\sum_{i=1}^N |BE_i^{\text{th}} - BE_i^{\text{exp}}|}{N}$$

## RESULTADOS

Se dividieron las 3556 masas del AME2020 en:

Training set = 3434 masas (AME2016)

Test set = 122 masas (2016-2020)

Fórmula para la desviación cuadrática media:

$$\sigma_{\text{rms}} = \sqrt{\sum_i^N \frac{(BE_i^{\text{th}} - BE_i^{\text{exp}})^2}{N}}$$

TABLE II. RMSDs for the training set (consisting of 3434 nuclei) and test set (consisting of 122 nuclei) achieved using different network structures.

Model	$\sigma_{\text{rms}}$ (MeV)		
	Training set	Test set	Entire set
ANN2	1.18	1.05	1.18
ANN4	0.55	0.63	0.55
ANN6	0.29	0.51	0.30
ANN7	0.19	0.34	0.20

TABLE III. Comparisons between ANN7 and the WS4 model [22], for nuclei with  $Z \geq 8$  and  $N \geq 8$  compiled in AME2020, i.e., 3336 nuclei and 120 nuclei contained in our training set and test set, respectively.

Model	$\sigma_{\text{rms}}$ (MeV)		
	Training set	Test set	Entire set
ANN7	0.15	0.34	0.16
WS4	0.42	1.30	0.47

-Diferencia entre ANN2 y ANN7 para el entire set.

-Disminución del 50% del RMSD en ANN4 para el entire set.

-La consideración de los Shell effects reduce RMSD hasta 0.3MeV.

-La consideración del ASY lo reduce a 0.2 MeV.

-ANN7 es mas preciso que WS4 para los 3 conjuntos de datos.

-sin los núcleos con Z y N menor a 8, la RMSD de ANN7 pasa de 0.20 a 0.16MeV.

-El RMSD aumenta del training set al test set, mas en WS4.

-ANN2 es malo en núcleos par-par.

En la ANN4 esto mejora.

-ANN4 tiene RMSD grande si Z o N es mágico. Si ambos lo son es aun mayor.

En ANN6 esto se reduce.

-La ANN7 mejora en los núcleos neutrónicamente enriquecidos respecto a ANN6 (disminuye el RMSD en el test set).

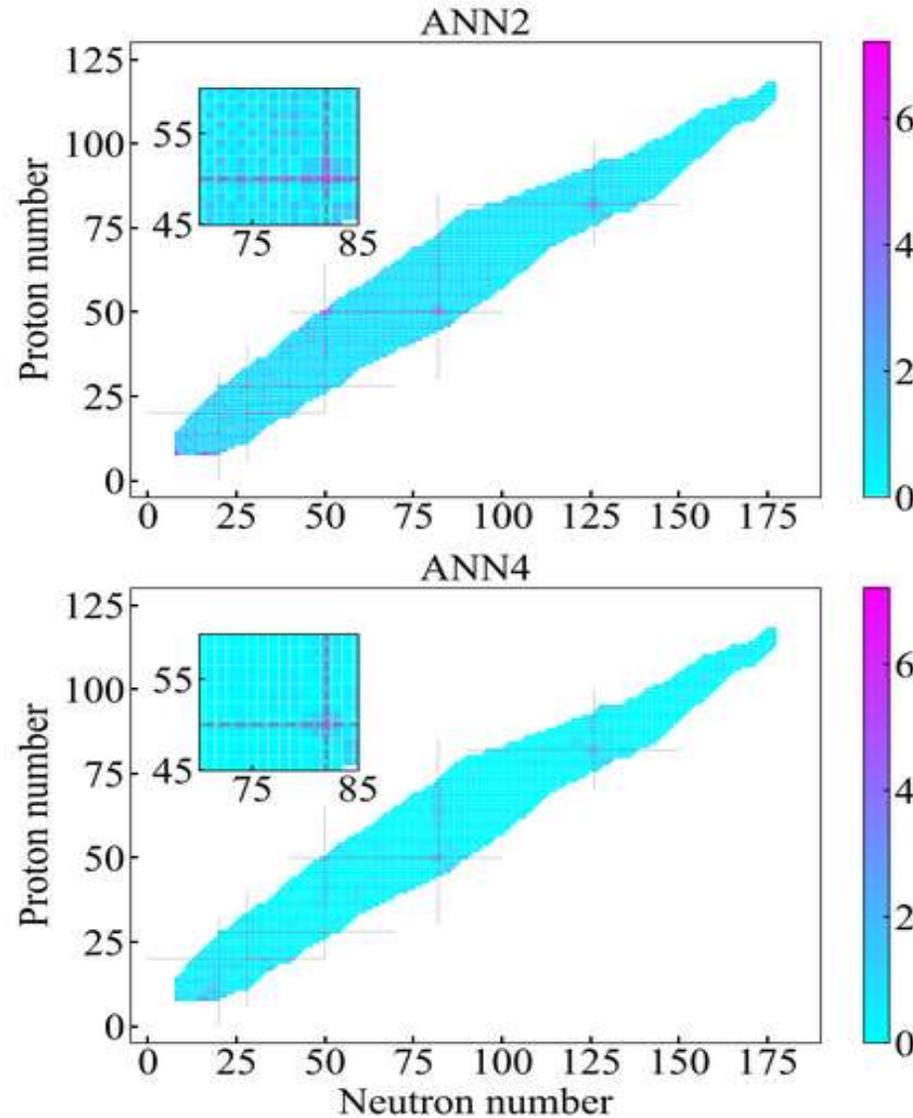


FIG. 2. Absolute deviations of the ANN2 and ANN4 predictions from the experimental binding energies. The gray lines denote the magic numbers.

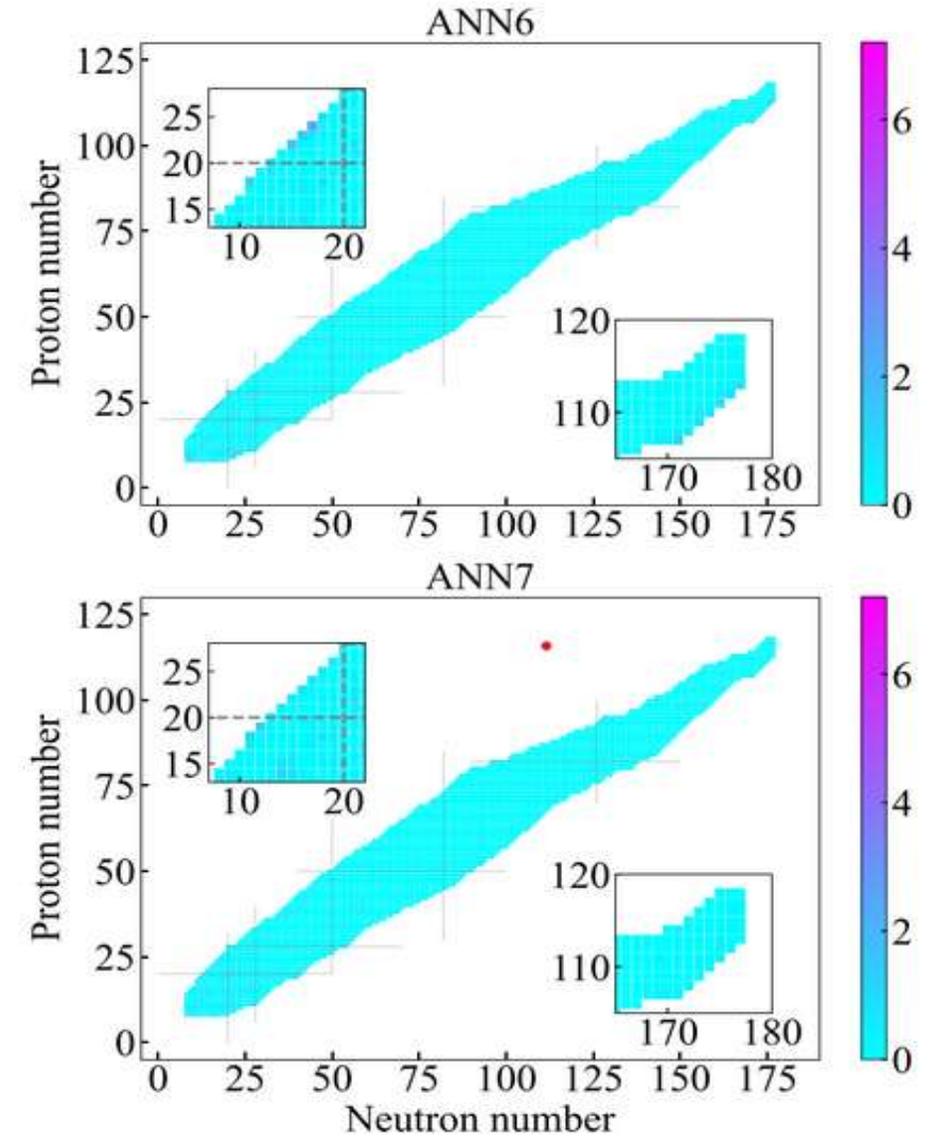


FIG. 3. Absolute deviations of the ANN6 and ANN7 predictions from the experimental binding energies. The gray lines denote the magic numbers.

# DISTRIBUCIÓN DEL RMSD EN FUNCIÓN DEL NÚMERO MÁSIICO

En promedio ANN7 tiene la desviación más baja.

Para núcleos livianos ( $A < 40$ ), ANN4 es el mejor.

Para núcleos con  $120 < A < 160$ , ANN6 supera por poco a ANN7.

ANN7 es mayor que ANN6 en núcleos livianos y pesados: el ASY ocupa un rol importante.

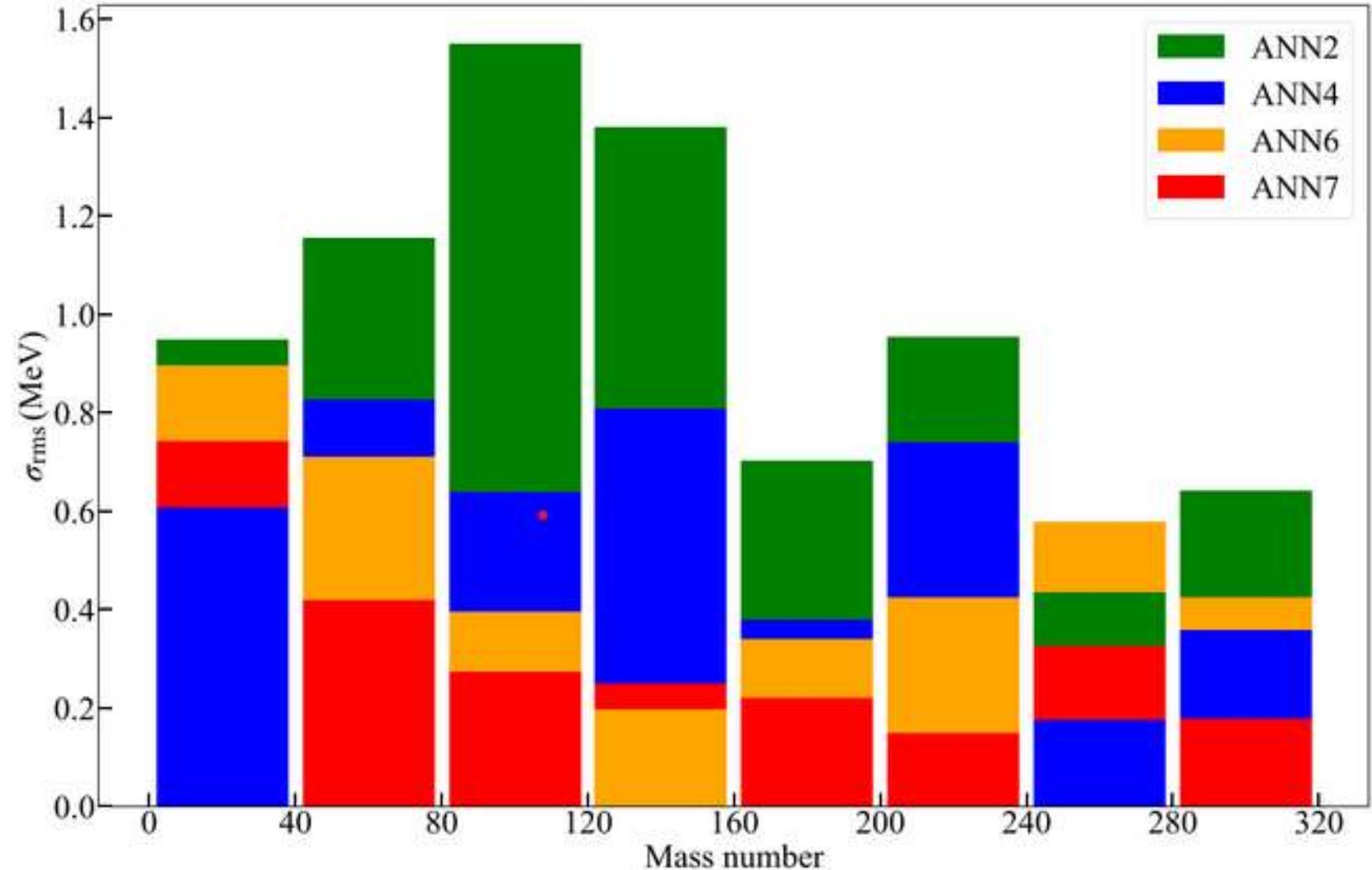


FIG. 4. Distributions of  $\sigma_{rms}$  between every 40 mass numbers in test20.

# COMPORTAMIENTO DE LAS ANN FUERA DE LA REGIÓN DE ESTABILIDAD

La ANN7 tiene las predicciones más precisas para núcleos fuera de la línea de estabilidad. (test set)

Para los núcleos con  $10 < N - Z < 20$  las predicciones de ANN6 son algo superiores.

No hay núcleos con  $50 < N - Z < 60$  en el test set.

Para la region donde la diferencia entre  $N$  y  $Z$  es mayor a 30 el RMSD de la ANN7 es menor o del orden de 0.2MeV.

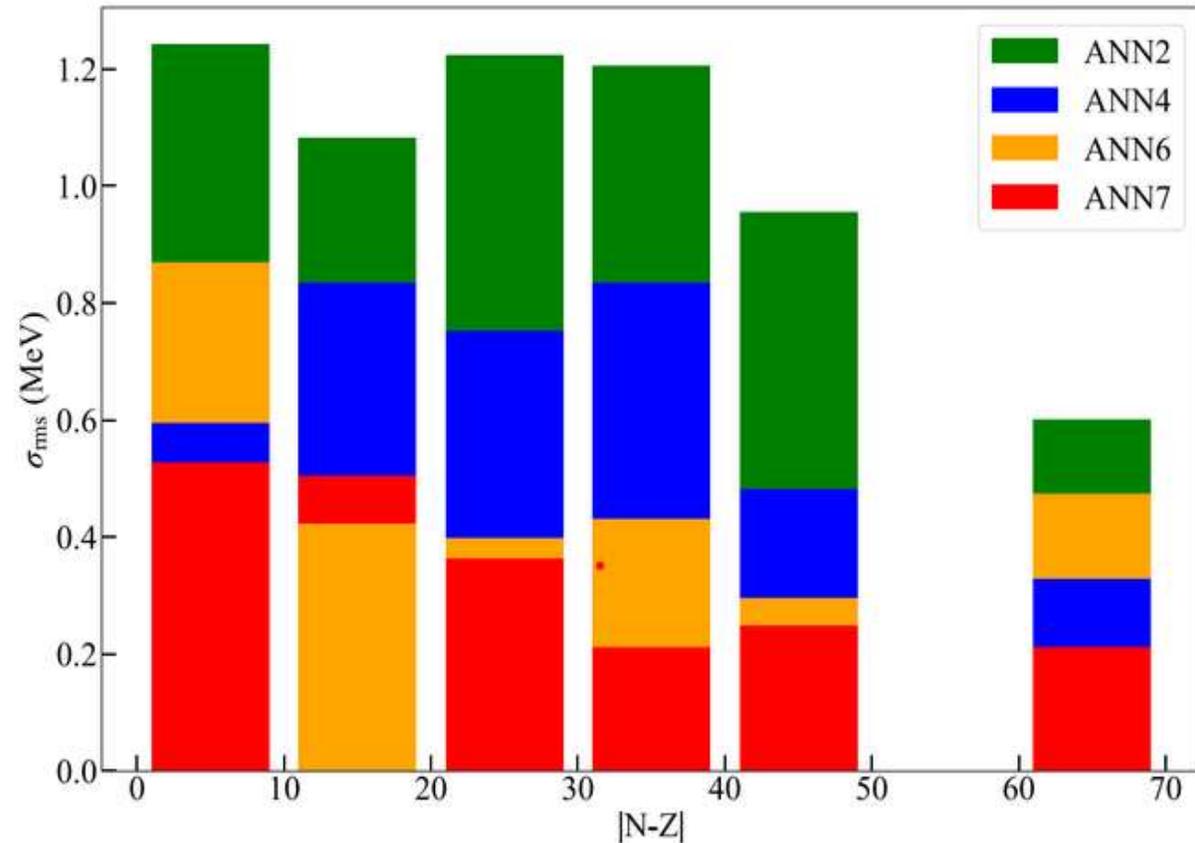
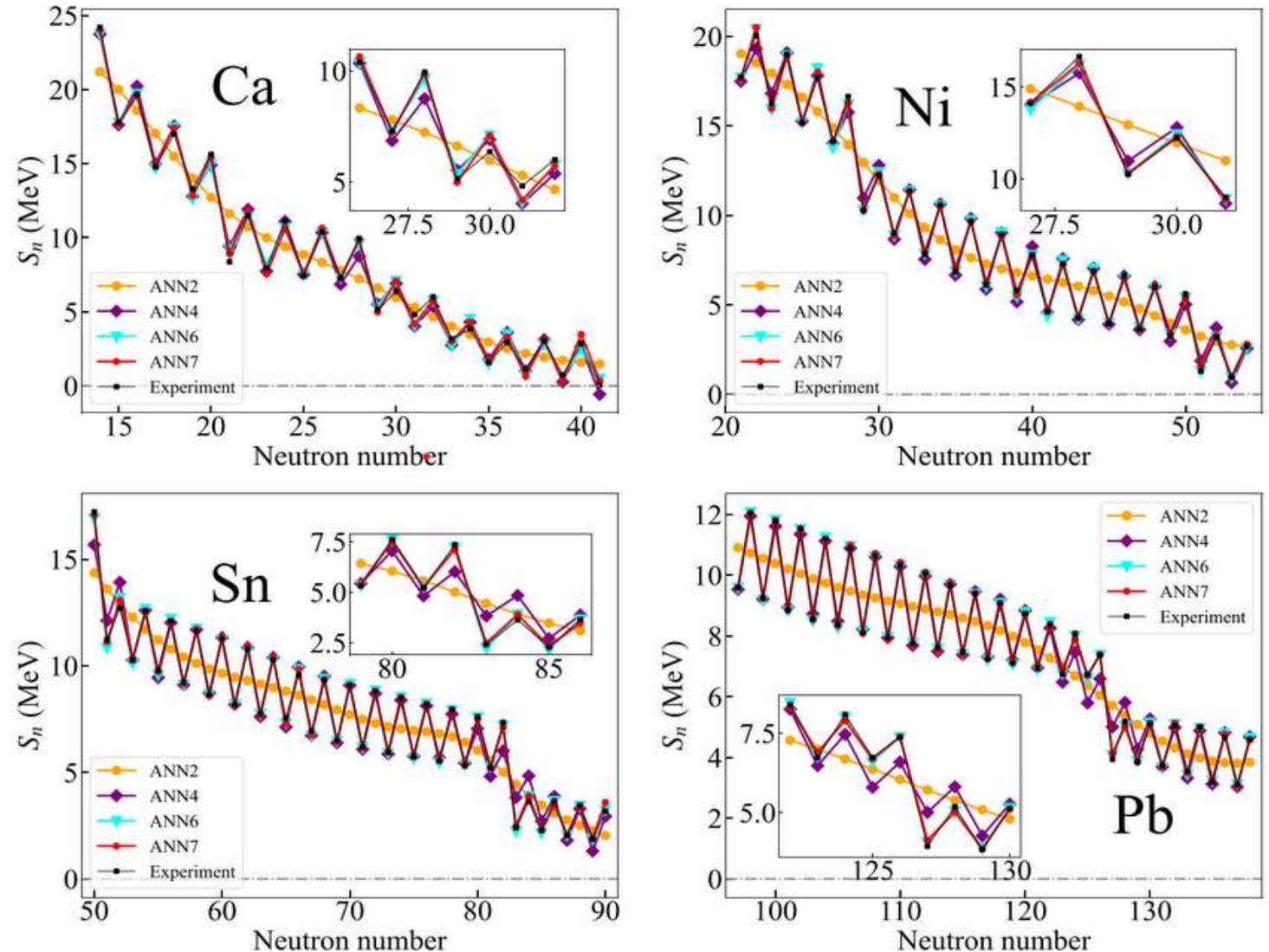


FIG. 5.  $\sigma_{rms}$  as functions of  $|N - Z|$ , which reflects the ability of ANNs to describe nuclei with large isospin asymmetries. Note that there is no nucleus in test20 with  $50 < |N - Z| < 60$ .

# CALCULO DE ENERGIA DE SEPARACION DE 1 NEUTRON CON LAS ANN.

La ANN2 no puede describir el escalonamiento par-impar de los núcleos.

La ANN4 corrige esto, pero su desviación es grande cerca de números mágicos.

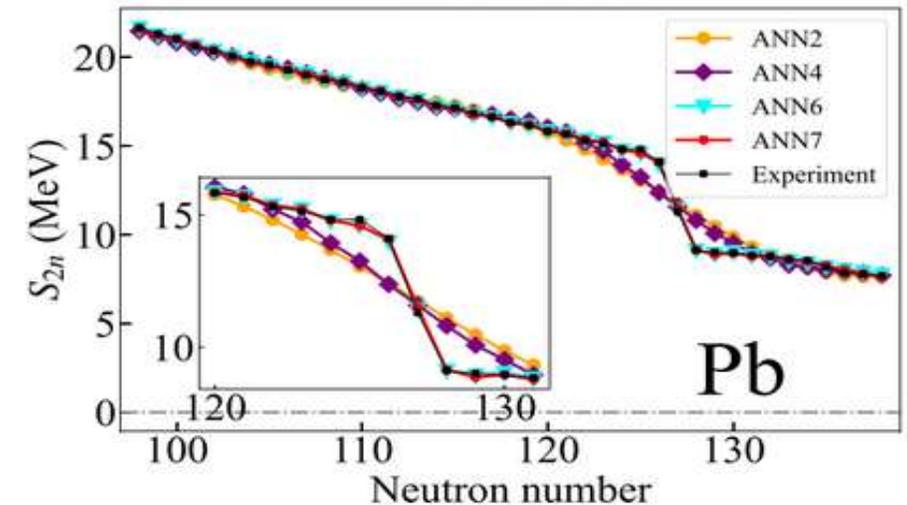
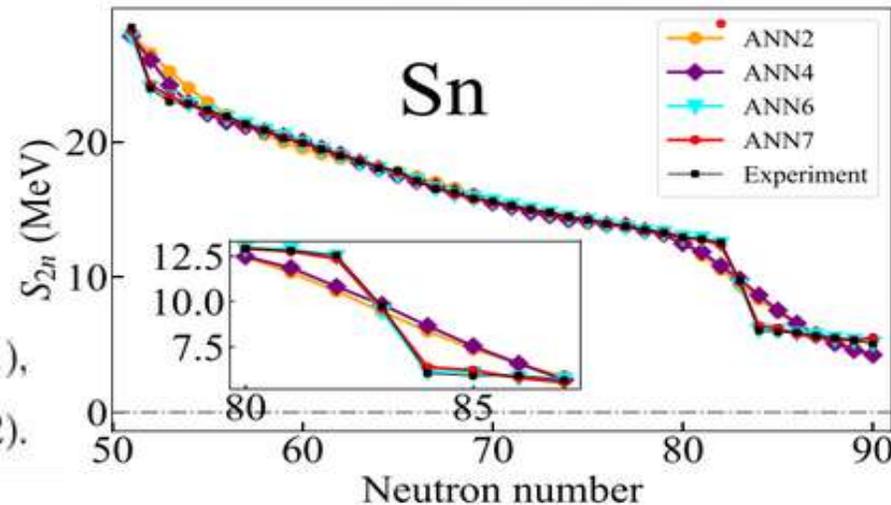
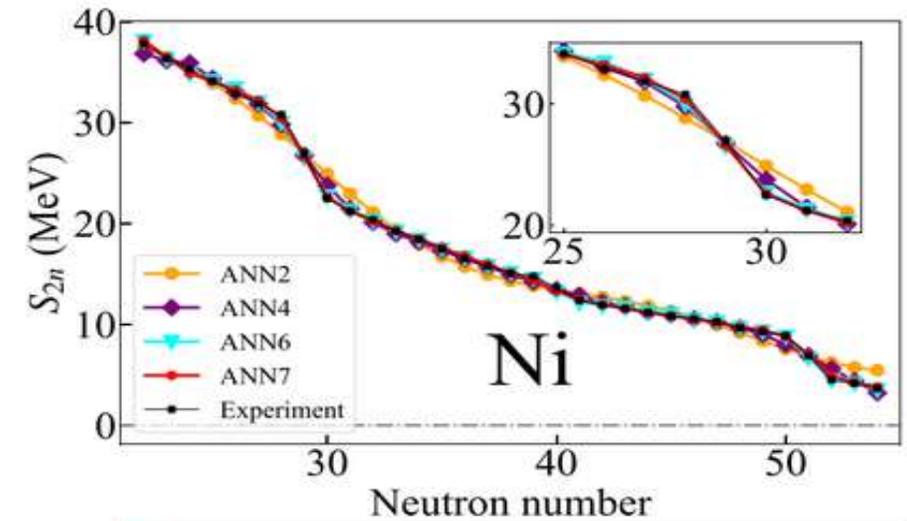
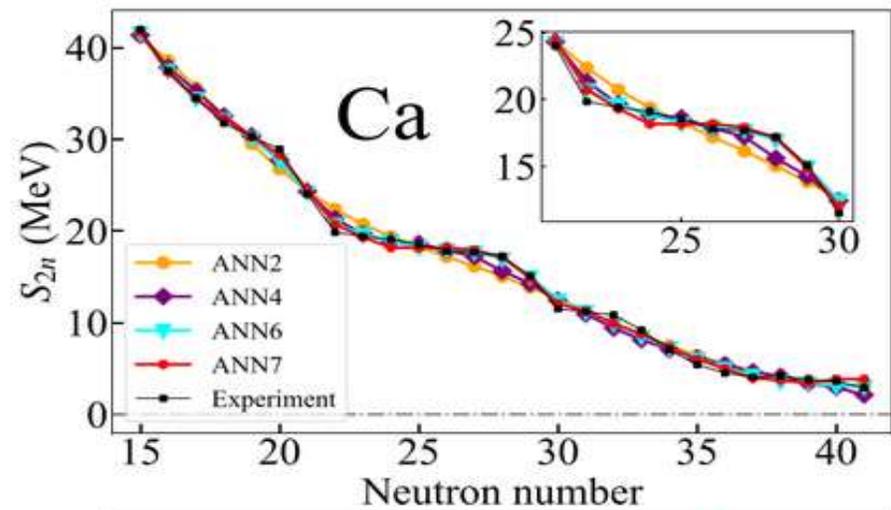


$$S_n(Z, N) = BE(Z, N) - BE(Z, N - 1),$$

$$S_{2n}(Z, N) = BE(Z, N) - BE(Z, N - 2).$$

FIG. 6. Experimental single neutron separation energies in comparison with the ANN predictions.

## CALCULO DE ENERGIA DE SEPARACION DE 2 NEUTRONES CON LAS ANN.



$$S_n(Z, N) = BE(Z, N) - BE(Z, N - 1),$$

$$S_{2n}(Z, N) = BE(Z, N) - BE(Z, N - 2).$$

FIG. 7. Experimental two-neutron separation energies in comparison with the ANN predictions.

# CANTIDAD DE NÚCLEOS CON DESVIACIONES MENORES A 0.7 MEV

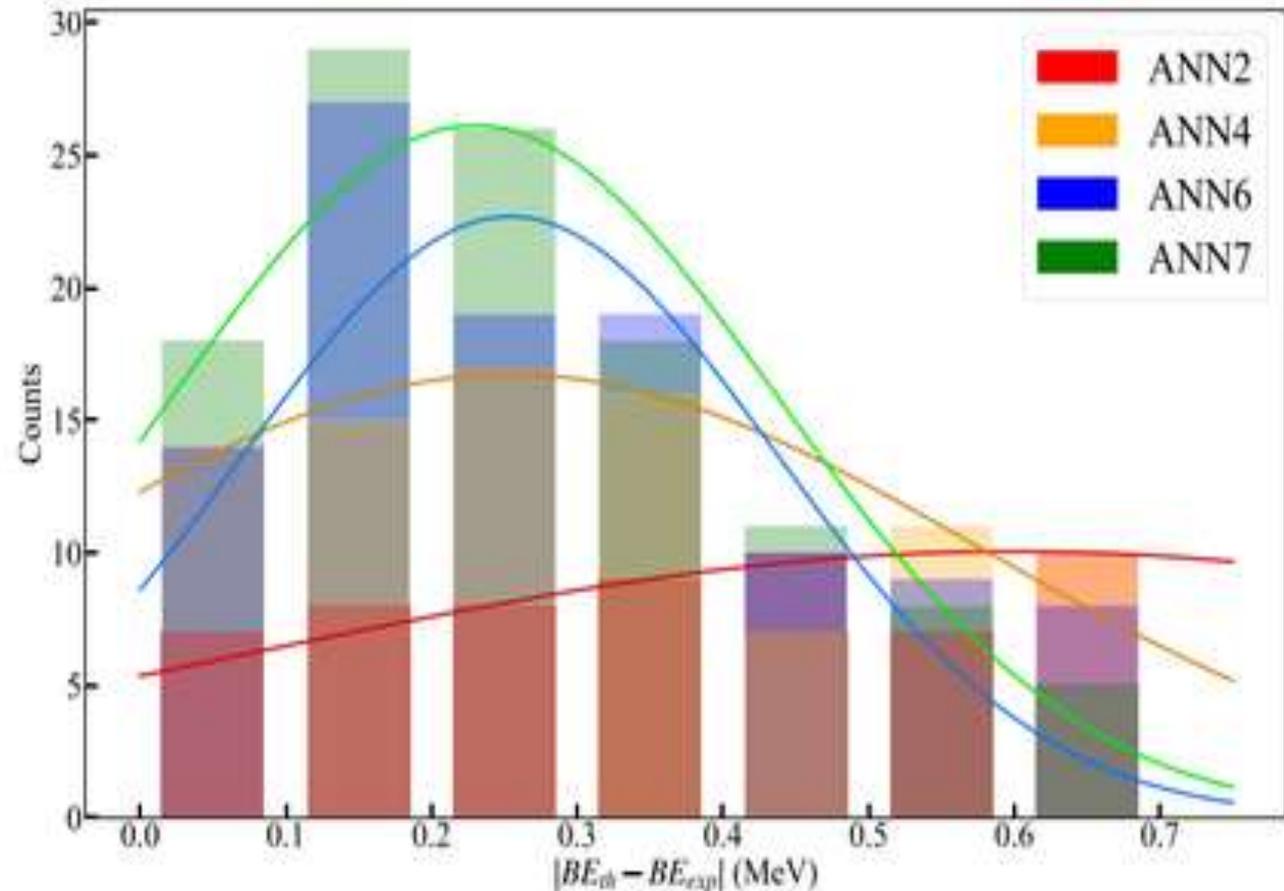
Se realizaron ajustes de los datos con distribuciones normales.

Porcentaje de núcleos en los que el RMSD es mayor a 0.7 MeV:

ANN2	51,6%
ANN4	26,2%
ANN6	13,1%
ANN7	5,7%

La media de la gaussiana indica la precisión de las mediciones.

La varianza indica el rango de desviación de las mismas.



## RESUMEN Y CONCLUSIONES.

- Se desarrolló una red neuronal artificial profunda con 7 inputs de entrada,  $Z$ ,  $N$ ,  $Z_{\text{eo}}$ ,  $N_{\text{eo}}$ ,  $\Delta Z$ ,  $\Delta N$  y  $ASY$ .
- Se calcularon las BEs de los núcleos del AME2020.
- Se destaca la importancia de los inputs ingeniosos cuando el data set es pequeño.
- Los métodos de aprendizaje automático son potentes para predecir BEs, con una precisión igual o mayor a muchos métodos teóricos convencionales.
- Por ejemplo, se pueden usar las predicciones para estudiar la nucleosíntesis del  $r$  - process o predecir las vidas medias de los posibles emisores de protones, lo cual resulta útil para planificar experimentos futuros.



FIN