

TERMOHIDRÁULICA Y NEUTRÓNICA

UTILIZACIÓN DE REDES NEURONALES PARA PREDECIR EL FLUJO DE CALOR CRÍTICO EN UN REACTOR

JAVIER RIVEROLA GURRUCHAGA
ENUSA INDUSTRIAS AVANZADAS, S.A.

En la industria nuclear de los reactores refrigerados por agua se considera una variedad de correlaciones estadísticas para la predicción del flujo calorífico basadas en resultados de ensayos con diferentes geometrías y diseños de combustible. Algunas son correlaciones públicas y otras son propiedad de los diferentes tecnólogos. Estas correlaciones son fiables y útiles, aunque algebraicamente complejas con poco intuitivas interdependencias entre los parámetros. En este trabajo se explora la utilización de redes neuronales como alternativa a las mencionadas correlaciones estadísticas para reactores PWR con diseños actuales de combustible sobre las mismas bases de datos experimentales. Además, se exploran diferentes funciones de activación, algoritmos de optimización, así como configuraciones con una o más capas ocultas. La red finalmente seleccionada predice el flujo calorífico de calor con una precisión significativamente mayor que las correlaciones estadísticas licenciadas. Se han comparado las distribuciones de medido-predicho, evidenciándose menor dispersión, así como un notable grado de validación, resultando en límites de diseño más reducidos que podrían proporcionar mayor margen de operación. Puede concluirse que las técnicas de inteligencia artificial tienen un potencial importante para abordar revisar los métodos de análisis de reactores nucleares bajo una nueva e interesante perspectiva.

UNA INTRODUCCION A LAS REDES NEURONALES

Inspirado en el funcionamiento de las neuronas biológicas, descubierto y descrito por el premio Nobel español Santiago Ramón y Cajal, el polifacético psicólogo y neurocientífico Frank Rosenblatt concibió en 1957 la primera neurona artificial o *perceptron* [1]. En su concepción más simple, el **perceptron** es una estructura lógica que pondera y adiciona varias señales de entrada, incluyendo un posible sesgo, y el resultado es a su vez procesado por una denominada función de activación, que aproxima el proceso en el que el axón de una neurona es activado o no. Algunas funciones de activación especiales (sigmoide, tanh, RELU,...), presentan unas interesantes propiedades matemáticas que hacen posible el mecanismo de retropropagación o revisión supervisada de pesos y sesgos durante el aprendizaje de la neurona artificial.

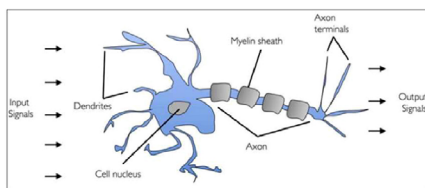
Siguiendo con el paralelismo entre inteligencia natural y artificial, una **red neuronal** artificial (ANN) es un conjunto de

perceptrones conectados entre sí de acuerdo a una estructura y jerarquía dadas, y son alimentadas por señales de entrada y resultan en variables de salida. Lo asombroso de las ANN es su capacidad de aprender. Esto es posible gracias a una revisión progresiva de los pesos y sesgos desconocidos en un algoritmo que minimiza el error, conocido como *retropropagación*, resuelto por Paul Werbos en 1975 [2].

En el ámbito de la Inteligencia Artificial, el diseño de redes neuronales ha evolucionado a una especialidad independiente, existiendo desde redes muy simples (*feed forward*, *deep forward*) hasta estructuras muy complejas (*deep convolutional*, *extreme learning machine*). En el momento de escribir este artículo, se ha llegado a un cerebro artificial con 16 millones de perceptrones, complejidad equivalente a un cerebro de rana. Pero sin necesidad de llegar a algo tan sofisticado, las ANN se han utilizado para resolver una amplia variedad de tareas, como la visión por computador y el reconocimiento de voz, y algoritmos de todo tipo, que son difíciles de resolver usando la programación ordinaria basado en reglas. No obstante, a pesar de su potencia y buen



Santiago Ramón y Cajal
(1852 – 1934)



Frank Rosenblatt
(1928–1971)

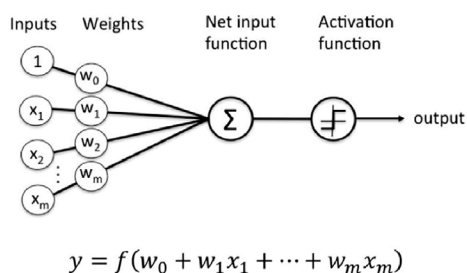


Figura 1. Santiago Ramón y Cajal (1852–1934) y Frank Rosenblatt (1928–1971).

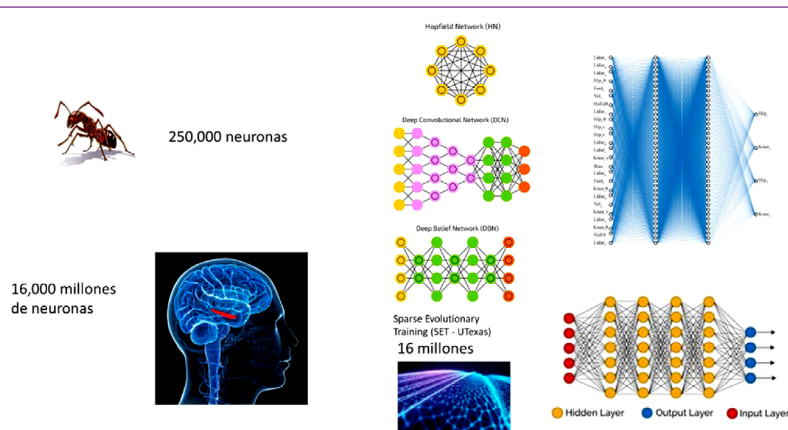


Figura 2. Cerebros biológicos y redes neuronales.

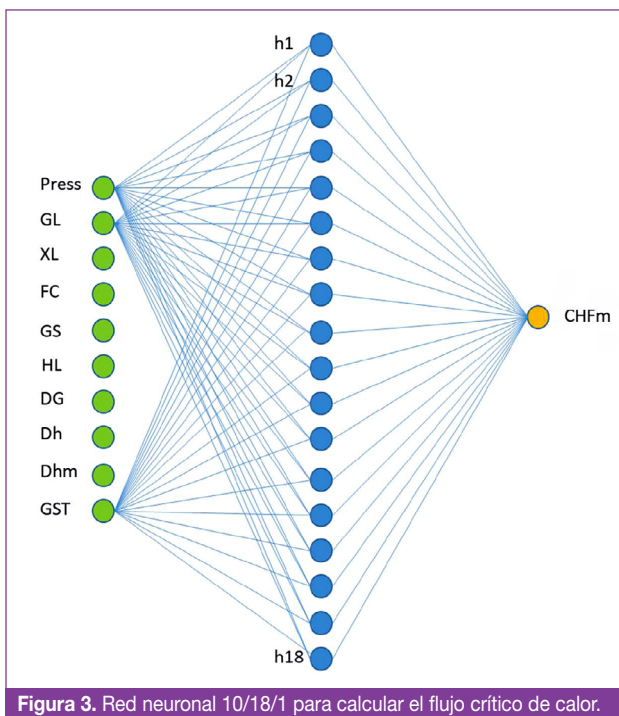


Figura 3. Red neuronal 10/18/1 para calcular el flujo crítico de calor.

comportamiento, una ANN no siempre es la solución ideal, y su idoneidad debe ser evaluada con cuidado en cada caso.

RED NEURONAL PARA PREDECIR EL FLUJO CRÍTICO DE CALOR

En el ejercicio que se presenta aquí, se ha partido de una base de datos de ensayos de flujo crítico de calor obtenidos en un lazo termohidráulico. En la sección de ensayo, se aloja un elemento combustible dotado de resistencias eléctricas para generar flujo calorífico y termopares para medir la distribución de temperaturas en la vaina y refrigerante, e instrumentación de medición del caudal y la presión del sistema. El ensayo de flujo calorífico crítico consiste en, partiendo de condiciones termohidráulicas estabilizadas, elevar progresivamente la potencia generada hasta obtener un súbito aumento de la temperatura de vaina en los termopares. Las condiciones locales del suceso (caudal másico local, calidad, entalpía, y flujo calorífico) se determinan analíticamente por medio de un modelo termohidráulico de subcanal del elemento combustible. Finalmente, el conjunto de condiciones locales y globales constituye la base de datos experimental y se toma como entrada para establecer las regresiones múltiples utilizadas en el ámbito del diseño y recogidas en la literatura pública o propietaria de los diferentes tecnologías.

Utilizando la librería Scikit-learn en la plataforma de programación PYTHON 3.7 junto con las librerías Numpy y Matplotlib entre otras, se ha diseñado un conjunto de redes

neuronales de aprendizaje supervisado. Consistentemente con la base de datos experimental utilizada, la capa de entrada incorpora cinco nodos para las condiciones termohidráulicas locales y otros cinco nodos más para contemplar la geometría del elemento combustible y posicionamiento de las rejillas mezcladoras. Para sortear el peligro de sobreajuste, y de acuerdo a las recomendaciones dadas en los grupos de usuarios, se ha considerado solamente una sola capa oculta, y limitado el número de perceptrones a aproximadamente el doble de los perceptrones de entrada. Para la salida, se toma únicamente un perceptron correspondiente el flujo calorífico crítico, q''_{cr} . Esta es una ANN sencilla, pero eficiente para los propósitos de este trabajo.

Por otra parte, en vez de utilizar todos los datos disponibles para realizar el ajuste, se ha reservado aproximadamente un 25 % para la validación de la red. Esta es una práctica necesaria no solo en los ajustes de una red neuronal, sino también en los ajustes estadísticos clásicos, algo que no siempre se ha seguido en el pasado.

Se ha probado con varias funciones de activación encontrando que la función rectificación o $Relu(x)$ da buen ajuste permitiendo que la retro propagación durante el aprendizaje proporciona un funcionamiento adecuado y ágil. También

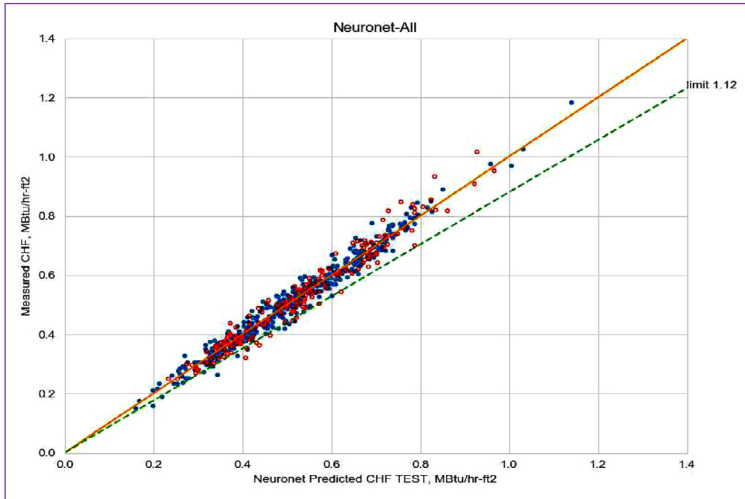


Figura 4. Valores predichos por la red neuronal frente a los medidos.

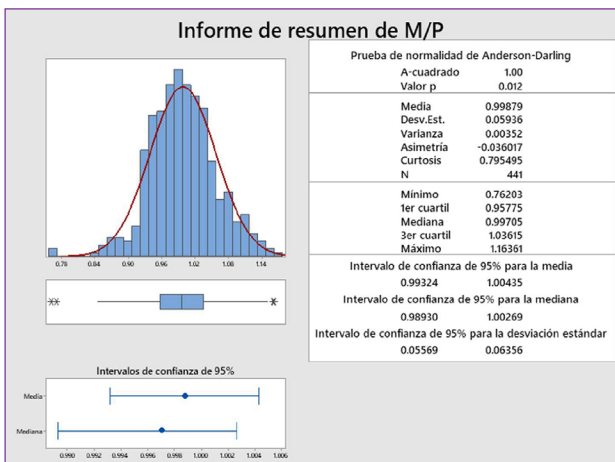


Figura 5. Distribución de valores M/P.

se ha escogido el algoritmo de optimización basado en el método de aproximación quasi-Newton (lbfgs). Otros parámetros para el optimizador como tolerancia y tasa de aprendizaje se han dejado en los valores por defecto, salvo el término de regularización o resolución del ajuste que se ha tomado 0.1, y una semilla para la aleatorización de búsquedas prefijada. La elección de estos hiperparámetros ha requerido una exploración de un espacio de múltiples posibilidades con varios óptimos locales, pero se ha conseguido una combinación que resulta en una incertidumbre reducida de las predicciones junto con un buen comportamiento en el subconjunto de las validaciones.

RESULTADOS Y DISCUSIÓN

La arquitectura de red descrita se ha aplicado a bases de datos de ensayos de flujo crítico de calor con combustible con rejillas mezcladoras representativo de un reactor PWR, en un amplio rango de presiones y caudales de operación.

La figura de mérito escogida para comparar el performance de la red neuronal es la razón entre el flujo de calor medido y el predicho (M/P) por la red neuronal así como el grado de validación de la red. En los casos analizados y como es de esperar en un ajuste, el valor de M/P es prácticamente la unidad (1000). No obstante, la dispersión de las predicciones resulta claramente inferior a la obtenida con los

métodos estadísticos clásicos, lo cual se traduce en una disminución de la incertidumbre de la red neuronal y del límite de aplicabilidad de la correlación para el cálculo del CLEN en aproximadamente un 6 %, que es un valor muy significativo.

Además, se ha visto también que las dispersiones individuales de M/P respecto de cada uno de los parámetros independientes son también menores y más uniformes que con las correlaciones clásicas, no observándose tendencias en forma de curva ni sesgos apreciables. Esto es debido a que las redes neuronales son tan eficientes en los extremos como en el centro de los intervalos y no existen zonas privilegiadas.

Ciertamente, una mejora del límite de correlación de un 6 % es muy significativo en términos de caudal, presión, potencia y Fdh, valores todos que intervienen en los límites térmicos del

núcleo de un reactor.

También representa una ganancia de margen importante en el cálculo del CLEN en transitorios, algo que en ocasiones es casi necesario.

Una cuestión de especial importancia es la validación de la red neuronal. Como se ha dicho anteriormente, un 25 % de los datos de los ensayos fue segregado de manera aleatoria y no ha intervenido en el entrenamiento de la red sino en la validación. Las predicciones de flujo calorífico para ese subconjunto de datos con la red neuronal ha sido satisfactoria, mostrando una estadística (media y sigma) parecida a las del subconjunto utilizado en el entrenamiento. Esto aporta gran confianza a la calidad del ajuste y puede considerarse que la ANN está debidamente validada y que no hay un sobreajuste apreciable.

Una vez calculada y validada la red neuronal, esta es aplicable para el mismo tipo de combustible dentro del rango experimental de cada uno de los parámetros. Para la red neuronal descrita, la expresión de cálculo del flujo crítico de calor es:

$$q''_{crit} = W_1^T f(W_0^T z + b_0) + b_1$$

$$\text{con } z = (x - \mu_{reg})/\sigma_{reg}$$

donde x es el vector de entradas [presión, caudal másico, calidad, geometría,...], μ_{reg} es el vector de medias de las variables de entrenamiento, σ_{reg} es el vector de desviaciones típicas de las variables de entrenamiento, z es el vector de entradas estandarizadas, W_0 es la matriz de pesos de capa oculta, W_1 es el vector de pesos de capa de salida, b_0 es el vector de sesgos de capa oculta, b_1 es el sesgo de capa de salida, f es la función de activación rectificador RELU(x) = max(0, x), q''_{crit} es el flujo calorífico crítico predicho por la red neuronal.

La complejidad de esta expresión es solamente aparente y es mucho más simple que las correlaciones predictivas clásicas y fácilmente codificable en un código termohidráulico.

CONCLUSIONES

Sobre las bases de datos experimentales de los ensayos de flujo crítico de calor se ha obtenido una red neuronal simple que predice el flujo crítico de calor con más precisión que las correlaciones típicamente utilizadas en la industria. Esta

red neuronal puede utilizarse de la misma manera que una correlación, pues se puede expresar fácilmente con una expresión analítica.

Esta mayor precisión se traduce en una menor incertidumbre y por tanto, en un menor límite de aplicabilidad de la correlación, casi un 6 % menor, lo cual es una mejora notable dentro del ámbito de las correlaciones de predicción de flujo calorífico.

El ejercicio aquí expuesto es solamente un ejemplo pero, dado que los métodos son enteramente aplicables a otras correlaciones de flujo calorífico y a otras magnitudes termohidráulicas, termomecánicas, etc., puede considerarse este resultado de un gran potencial.

Queda pendiente el estudio de otras arquitecturas y algoritmos de aprendizaje de máquina.

REFERENCIAS

- [1]. Rosenblatt, Frank (1957). "The Perceptron—a perceiving and recognizing automaton". Report 85- 460-1. Cornell Aeronautical Laboratory.
- [2]. Paul J. Werbos. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis, Harvard University, 1974■