

Empleo de redes neuronales para determinar la calidad de secuencias aleatorias

Juliana Bourdieu, Myriam Nonaka, Mónica Agüero, Marcelo Kovalsky, and
Alejandro Hnilo

Laboratorio de Láseres Sólidos, Centro de Investigaciones en Láseres y Aplicaciones
(CEILAP), UNIDEF, Instituto de Investigaciones Científicas y Técnicas para la
Defensa (CITEDEF), Villa Martelli, Buenos Aires, Argentina.
`{jbourdieu,mnonaka,magüero,mkovalsky,ahnilo}@citedef.gob.ar`

Abstract. La calidad de la aleatoriedad es una característica central en la seguridad de las claves criptográficas. Los test estadísticos (NIST, DIE HARD, etc.) son sólo una primera aproximación al tema. En este trabajo se comparan dos secuencias que pasan todos los test estadísticos de aleatoriedad, una pseudorandom generada por computadora y otra de origen físico, proveniente de los *outcomes* de un experimento de pares de fotones entrelazados en polarización, que según una interpretación de la Mecánica Cuántica son las únicas “verdaderamente aleatorias”. Para evaluar su desempeño, empleamos una red neuronal ESN (Echo State Network) para predecir series temporales de trayectorias del atractor de Lorenz, cuando el reservorio de la red es construido a partir de cada una de estas secuencias. El error en la predicción en cada caso es el parámetro empleado para determinar la calidad de cada secuencia aleatoria.

Keywords: redes neuronales, computación por reservorio, aleatoriedad.

1 Introducción

Las redes neuronales artificiales son algoritmos que simulan el mecanismo de aprendizaje en los organismos biológicos. El sistema nervioso humano contiene células, conocidas como neuronas. Las neuronas están conectadas entre sí mediante axones y dendritas, y las regiones de conexión entre axones y dendritas se denominan sinapsis. La intensidad de las conexiones sinápticas a menudo cambia en respuesta a estímulos externos. Es a través de este cambio que ocurre el aprendizaje en los organismos vivos. Las redes neuronales artificiales contienen unidades de cálculo denominadas neuronas que están conectadas entre sí mediante pesos, que cumplen el mismo rol que las intensidades de las conexiones sinápticas en los organismos biológicos. Cada entrada a una neurona se escala con un peso, que afecta la función calculada en esa unidad. Una red neuronal artificial calcula una función de las entradas propagando los valores calculados desde las neuronas de entrada hacia la(s) neurona(s) de salida y utilizando los pesos como parámetros intermedios. El aprendizaje ocurre al cambiar los pesos que conectan las neuronas. Entre los muchos tipos de implementación de redes neuronales empleamos la técnica conocida como computación por reservorio

(reservoir computing). Reservoir Computing (RC) se basa en la idea de utilizar una red neuronal recurrente (denominada “reservorio”) para transformar datos de entrada en un espacio de alta dimensionalidad, y luego aprender a partir de esas representaciones usando un modelo de salida simple. La ventaja de RC es que el entrenamiento se realiza solo en una parte de la red, lo que simplifica el proceso y permite abordar problemas complejos de manera más eficiente.

Echo State Network (ESN) es un tipo de red neuronal recurrente (RNN) que se encuadra dentro de las técnicas de RC. Una ESN se compone de tres capas: una capa de entrada, una capa oculta conocida como reservorio y una capa de salida. Las neuronas de la capa de entrada y las de la capa del reservorio están completamente conectadas con pesos fijos, asignados aleatoriamente. La señal de entrada induce una respuesta no lineal en cada neurona del reservorio y, mediante un mecanismo de lectura con un algoritmo de aprendizaje simple, es posible aplicarlo en la predicción de series temporales. La implementación de este tipo de red neuronal típicamente utiliza números pseudoaleatorios [1] generados por algún algoritmo. En este trabajo evaluamos cómo influye la aleatoriedad en la calidad de la predicción de la ESN. Existen dos formas de generar secuencias aleatorias. Una, llamada pseudorandom, que se basa en algoritmos deterministas generados por computadora que, a partir de una semilla, producen secuencias de períodos extremadamente largos. La otra es a partir de un fenómeno físico. De particular interés son los cuánticos, ya que son considerados por algunas corrientes de pensamiento como los que producen “verdadera aleatoriedad” [4].

Para responder a la pregunta: ¿hay alguna diferencia al usar números generados por un proceso verdaderamente aleatorio en lugar de aquellos generados por un algoritmo de computadora? compararemos los resultados de predecir con una ESN y estas secuencias aleatorias, las componentes x , y , z , de las ecuaciones de Lorenz en una condición caótica.

2 Resultados

Generamos una serie temporal a partir de las ecuaciones de Lorentz para los ejes x , y , z durante 10149 pasos de tiempo, usando el algoritmo proporcionado por I. Moiseev et al. [2] con los clásicos parámetros que producen caos: $\rho = 28$, $\sigma = 10$ y $\beta = 8/3$. Parte de esta serie temporal se emplea para “entrenar” la red neuronal y otra para comparar con las predicciones.

En cuanto a la matriz aleatoria para crear el reservorio dinámico empleamos para los números aleatorios físicos una secuencia random proveniente de las salidas de un experimento de Bell [3] que genera pares de fotones entrelazados en polarización (números aleatorios cuánticos, QRN) mientras que los números aleatorios pseudorandom (PRN) son generados por el software MATLAB a través de los algoritmos Marsenne Twister, SIMD-Oriented Fast Mersenne Twister, Combined Multiple Recursive, Multiplicative Lagged Fibonacci, Philox y Threefry. En la tabla 1 se observa el resumen de los errores cuadráticos medios (RMSE) entre los valores predichos por la ESN con QRN y PRN para cada coordenada.

	QRG	Marsenne Twister	SIMD	Combined Multiple Recursive	Fibonacci	Philox	Threefry
RMSE_x	1.2462	3.2719	0.6402	0.6271	2.0463	0.47379	1.1563
RMSE_y	1.0534	5.1009	6.4139	6.5547	5.9915	5.5253	7.0076
RMSE_z	27.2126	27.5707	N.A.	27.1077	N.A.	27.1404	27.4575

Table 1. Error cuadrático medio (RMSE) adimensional en las coordenadas x , y y z entre la ecuación de Lorenz y los valores predichos (350 pasos) por la ESN con cada uno de las secuencias aleatorias, una física (QRG) y las restantes 6 generadas por algoritmos de computadora (PRN). N.A.: No Aplica.

3 Conclusiones

En primer término vemos que los errores obtenidos son diferentes, es decir hay una dependencia con el tipo de aleatoriedad empleada. No hay una clara ventaja para ninguna en particular. La predicción que empleó la matriz con aleatoriedad física, solamente resultó ser la mejor en la coordenada y , en la x resultó en quinto lugar y en z todas las predicciones tienen un error muy similar. Entonces, el esfuerzo que implica la generación de secuencias aleatorias provenientes de un sistema cuántico no parece recompensado con una notable mejora en la calidad de las predicciones. Si es de destacar el muy buen desempeño de la ESN en general, si tenemos en cuenta que el horizonte de predicción teórico que surge del cálculo de los exponentes de Lyapunov da un valor de 38 pasos y hemos podido predecir hasta 350 pasos.

Referencias

1. Matlab, “Control the random number generator”, último acceso: 31 de enero 2024. [Online]. Disponible en: <https://la.mathworks.com/help/matlab/ref/rng.html>
2. Matlab, “Lorenz attaractor plot”, último acceso: 31 de enero 2024. [Online]. Disponible en: <https://la.mathworks.com/matlabcentral/fileexchange/30066-lorenz-attaractor-plot>
3. M. Nonaka, M. Agüero, M. Kovalsky, and A. Hnilo, “Testing randomness of series generated in an optical bell’s experiment”, *Applied Optics* **62**, 3105–3111, (2023).
4. A. Acin, L. Masanes, “Certified randomness in Quantum Mechanics”, *Nature* **540**, 213–219, (2016).