

Reporte práctica doce: Redes neuronales

José Anastacio Hernández Saldaña

Posgrado de Ingeniería de Sistemas

1186622

jose.hernandezsld@uanl.edu.mx

30 de octubre de 2017

Resumen

Este es un reporte sobre la práctica doce con respecto al tema de redes neuronales que se realizó en la clase de Simulación de sistemas, cómputo paralelo en R.

1. Tarea: Paralelizar la red neuronal

Las redes neuronales son un modelo computacional de la simulación de una neurona en un cerebro biológico. Las neuronas responden a estímulos y si la cantidad de estímulos pasa de cierto umbral, la neurona se activa generando una respuesta o salida. Las neuronas artificiales se conectan a entradas, donde cada entrada tiene un valor de peso de importancia y de acuerdo a una función de activación que considera las entradas, la neurona artificial se activa y genera una respuesta. De esta manera las neuronas artificiales pueden reaccionar a patrones, donde si se ajustan los pesos de las entradas las neuronas artificiales pueden simular aprendizaje.

Para esta práctica se paralelizará el código de la página del curso donde se busca que una red neuronal aprenda los dígitos del cero al nueve dibujados en una cuadrícula de 5×3 , se cuenta con el modelo de los dígitos, donde cada cuadro está activado si tiene un valor verdadero, y falso si está desactivado. Para el entrenamiento y prueba de la red se tiene un modelo de cuadrícula que marca las probabilidades de que el cuadro esté activado o no, de acuerdo a la posible escritura del dígito. Se tienen tres posibles valores, la probabilidad de que un cuadro blanco esté activado es de $PB = 0.002$, la probabilidad de que un cuadro gris se active es de $PG = 0.92$, y la probabilidad de que un cuadro negro se active es de $PN = 0.995$. En la figura 1 podemos ver un ejemplo del modelado.

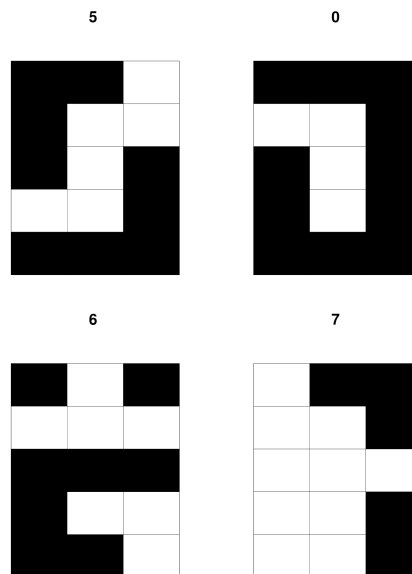


Figura 1: Ejemplos de modelado

1.1. Diseño del Experimento

Para el experimento se buscó paralelizar el código de la red neuronal, el cual consta de dos fases, una fase de entrenamiento y otra de prueba, donde se espera que la red neuronal categorice adecuadamente los números que se le presentan.

Las redes neuronales tienen una dependencia en su fase de entrenamiento, ya que cada iteración depende del ajuste hecho a la neurona en iteraciones anteriores, por lo que no se podrá paralelizar este proceso y solo se paralelizó la fase de prueba.

Para la experimentación, se utilizó una computadora con las siguientes especificaciones, Procesador Intel Core i7-4790 CPU @ 3.6GHz \times 8 y memoria RAM de 24 GB, utilizando solamente los cuatro núcleos físicos disponibles.

Se utilizaron los siguientes parámetros para la experimentación, se realizaron 30 réplicas y se utilizaron tamaños de prueba de 300, 600 y 1200 iteraciones para ver la interacción del tamaño con los tiempos de ejecución, los códigos están disponibles en el repositorio git del curso.

1.2. Resultados

Al revisar los resultados de la experimentación, que están graficados en la figura 2, se puede ver los resultados del porcentaje de aciertos para cada tipo de procedimiento y cómo aunque se aumente el tamaño de prueba, esto no afecta la clasificación de los dígitos por parte de la red neuronal.

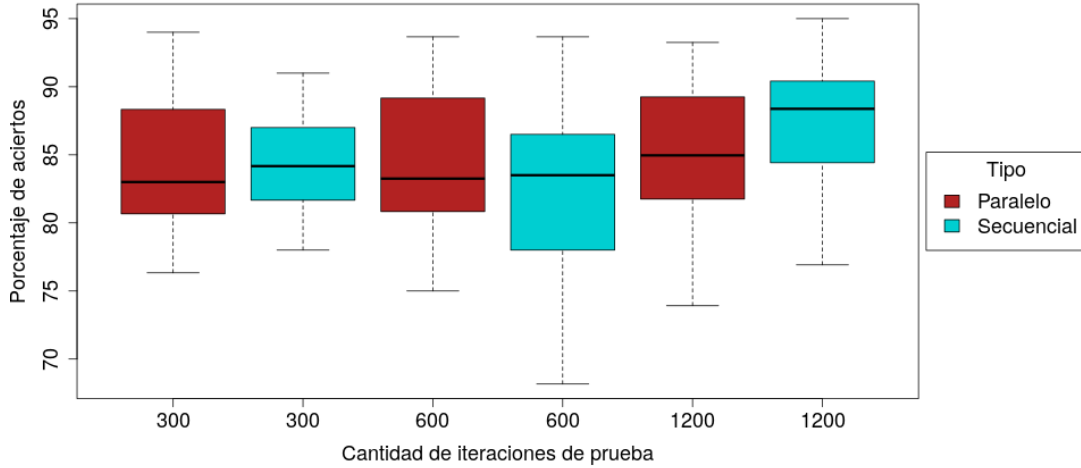


Figura 2: Gráfica del porcentaje de aciertos con los tipos y tamaños de muestras

En cambio, para el resultado de tiempo de la figura 3 podemos apreciar que sí hay una mejora conforme aumenta el tamaño de la muestra, incluso se ve que en la muestra de tamaño de 300 la versión secuencial es más rápida que la versión paralela y conforme aumenta el tamaño, la versión paralela tarda menos en clasificar los dígitos.

Estos resultados muestran el beneficio de la paralelización en el tiempo de ejecución de la red neuronal, sin que esto afecte el desempeño de la cantidad de aciertos de la red neuronal.

2. Extra Uno: Desempeño de la red en función de las probabilidades del modelo de generación

Ahora se estudiará el efecto de las probabilidades asignadas en el modelo de representación de los dígitos en función del porcentaje de aciertos que la red registra.

2.1. Diseño del Experimento

Teniendo las 3 Probabilidades PN , PG , y PB se dividirán en 4 factores que corresponden a un porcentaje del valor inicial sugerido en la práctica, esto para ver las interacciones que hay entre las variaciones. Los factores son los siguientes:

N con un 0% del porcentaje original a que el cuadro se active.

B con un 33.33% del porcentaje original a que el cuadro se active.

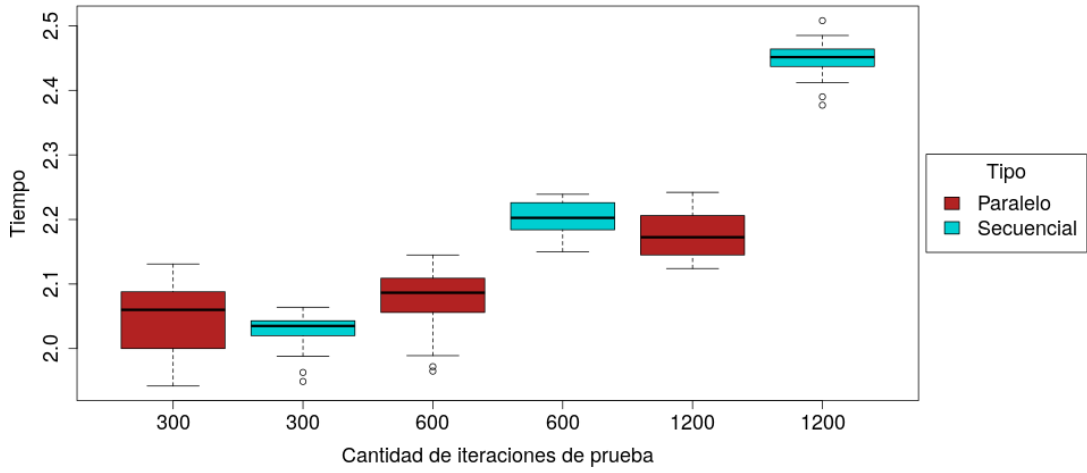


Figura 3: Gráfica del tiempo de ejecución con los tipos y tamaños de muestras

M con un 66.66 % del porcentaje original a que el cuadro se active.

A con un 100 % del porcentaje original a que el cuadro se active.

Utilizando estos factores se podrá apreciar el desempeño de la red neuronal cuando los datos de entrenamiento y prueba se muestran de manera parcial, esto dificultará la clasificación de los dígitos.

Los parámetros utilizados para el experimento fueron los cuatro factores en las tres probabilidades con 30 réplicas cada uno.

2.2. Resultados

El resultado obtenido de la experimentación se puede ver en las figuras 4, 5, 6, y 7 donde se puede ver para cada valor de PN el comportamiento de PG y PB con sus interacciones.

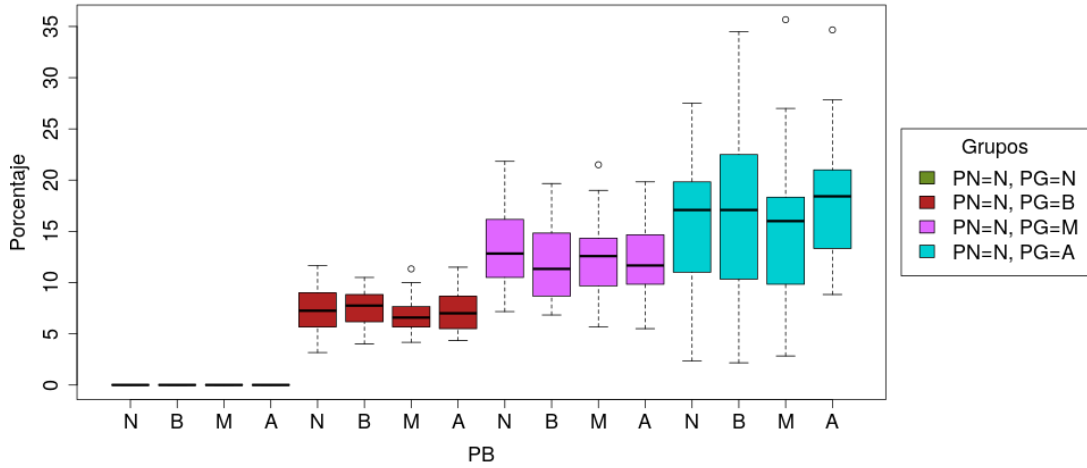


Figura 4: Interacciones de las probabilidades cuando $PN = N$

En la figura 4 están los resultados obtenidos al tener $PN = N$ con el incremento generado al subir PG y PB . Con estos valores de probabilidad donde las casillas negras están ausentes, vemos que la probabilidad de aciertos es baja, ya que es poca la información con la que cuenta la red para entrenarse y clasificar.

En la figura 5 están los resultados obtenidos al tener $PN = B$ con el incremento generado al subir PG y PB . Con estos valores de probabilidad donde las casillas negras tienen baja probabilidad de

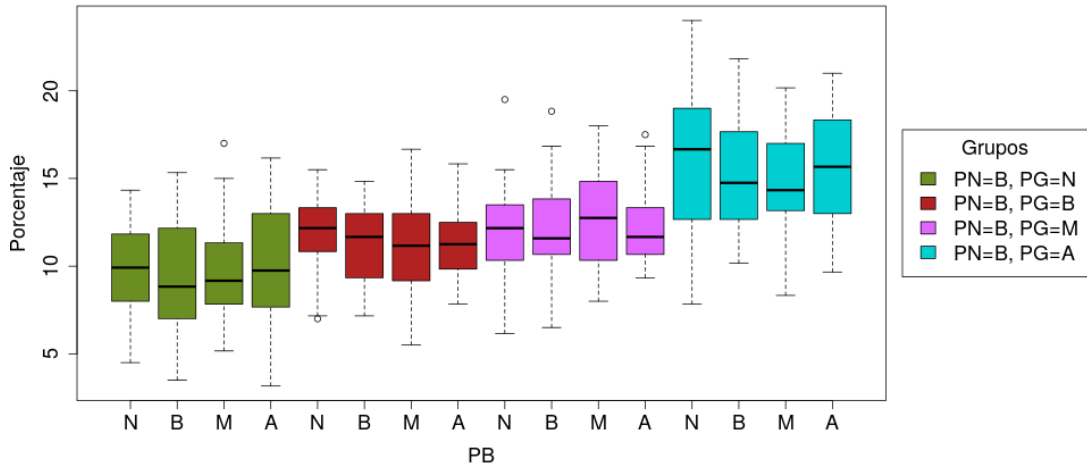


Figura 5: Interacciones de las probabilidades cuando $PN = B$

aparecer, la probabilidad de aciertos se incrementa con relación al caso anterior, aunque sigue siendo una probabilidad de acierto baja.

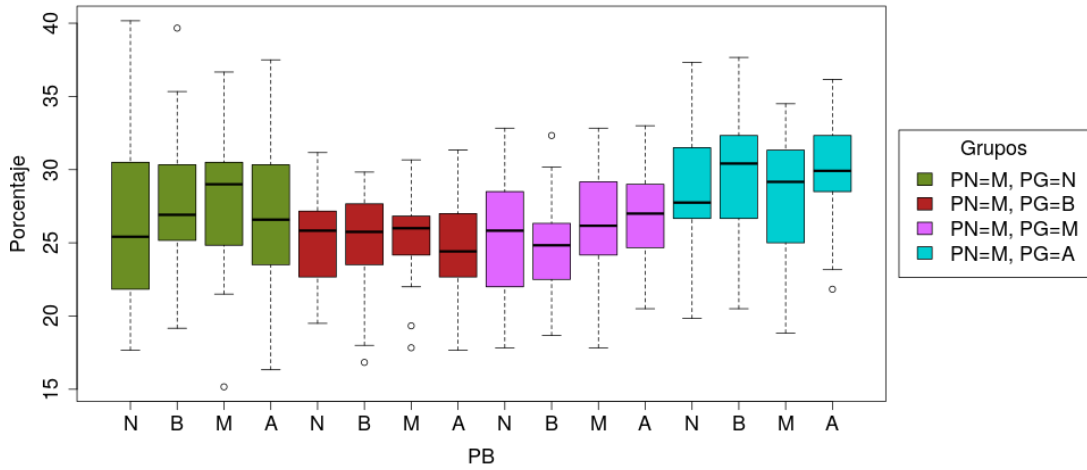


Figura 6: Interacciones de las probabilidades cuando $PN = M$

En la figura 6 están los resultados obtenidos al tener $PN = M$ con el incremento generado al subir PG y PB . Con estos valores de probabilidad donde las casillas negras tienen una probabilidad de aparecer media, la probabilidad de aciertos se incrementa también con relación al caso anterior, pero sigue siendo una probabilidad de acierto baja.

En la figura 7 están los resultados obtenidos al tener $PN = A$ con el incremento generado al subir PG y PB . Con estos valores de probabilidad donde las casillas negras tienen una alta probabilidad de aparecer, la probabilidad de aciertos se incrementó considerablemente con relación al caso anterior, marcando una transición de fase, ya que ahora la red neuronal puede clasificar mejor. El caso donde es muy notable, es cuando $PN = A$ con $PG = N$ ya que al estar casi siempre presentes solo los dígitos negros, pueden ser fácilmente identificados.

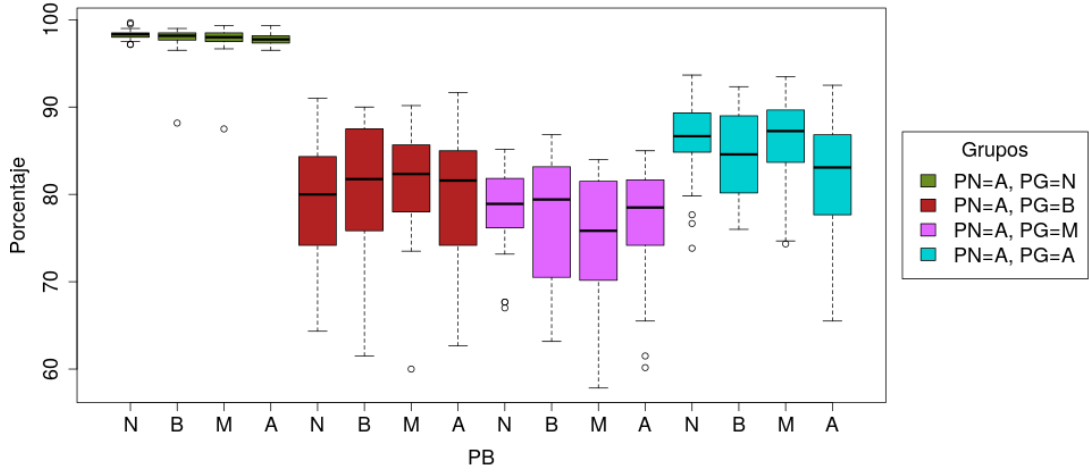


Figura 7: Interacciones de las probabilidades cuando $PN = A$

3. Extra Dos: Desempeño de la red con dígitos extendidos

Ahora se estudiará el desempeño de la red neuronal cuando se agregan más dígitos. Además de los dígitos del cero al nueve, se incluirán 12 más. Esto para apreciar el comportamiento de la red cuando se tienen más valores de entrada.

3.1. Diseño del Experimento

Para el experimento se agregaron las primeras 12 letras del alfabeto como se ve en la figura 8, este incremento hizo necesario agregar una neurona más, ya que cada neurona identifica el valor que corresponde a una posición de la representación binaria del número. Con cinco neuronas se puede categorizar hasta 2^5 valores.

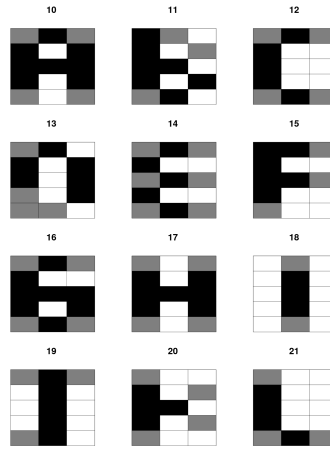


Figura 8: Caracteres agregados

Con el objetivo de ver el comportamiento de la red al subir la cantidad de caracteres, se incrementó de tres en tres el tamaño hasta llegar a los 22 caracteres. Se espera que el porcentaje de aciertos disminuya, por lo que también se probó incrementando la cantidad de iteraciones de la fase de entrenamiento con los valores de 5,000, 10,000, y 15,000. Como en los experimentos anteriores, se realizaron 30 réplicas por experimento.

3.2. Resultados

En la gráfica de la figura9, podemos ver cómo disminuyó el desempeño de la red conforme aumentó la cantidad de caracteres como se esperaba, pero es interesante ver que aunque se aumentó la cantidad de iteraciones de entrenamiento, no se ve una diferencia significativa en el desempeño de la red.

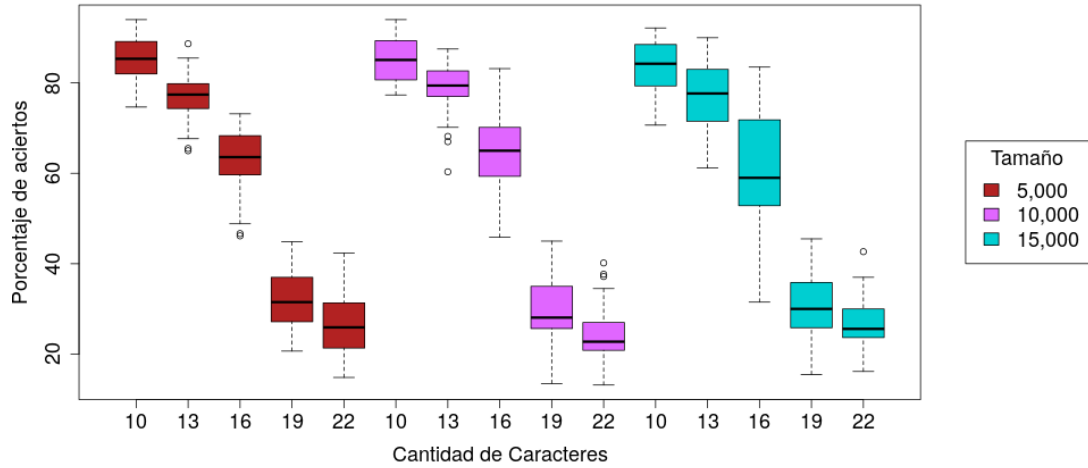


Figura 9: Desempeño de la red neuronal al incrementar la cantidad de caracteres

Esto es porque la cantidad de neuronas para la cantidad de dígitos posibles es mayor y hay más similitudes entre los dígitos.

4. Conclusiones

Se lograron observar las diferencias entre el desempeño de la red neuronal paralelizada y secuencial, además se mostró cómo afectan las formas de los dígitos en el resultado de la red, al variar las probabilidades que afectan a los cuadros. Por último, se vio cómo el desempeño de la red es afectado con respecto a la cantidad de datos en la entrada y la cantidad de iteraciones de entrenamiento.