

**Laboratorio No.1**  
*Redes Neuronales Profundas*

1. El ancho (tamaño de la capa escondida) del algoritmo. Intenten con un tamaño de 200.
  - a. ¿Cómo cambia la precisión de validación del modelo?
    - i. Si, la precisión del modelo aumenta un 0.0124 en comparación a las 50 épocas con las que se trabajó ya que estas en un inicio dieron 96.57% de precisión mientras que ahora por las 150 nuevas capas dio 97.81%
  - b. ¿Cuánto tiempo tardó el algoritmo en entrenar?
    - i. Inicialmente el algoritmo se tomó 12 segundos para entrenar toda la data y este subió un poco al momento de cambiarle la capa escondida a 17 segundos esto se puede deber a que al aumentar las capas permite procesar representaciones más complejas.
  - c. ¿Puede encontrar un tamaño de capa escondida que funcione mejor?
    - i. Si, al momento de ir probando distintos números de capas puede observar que obtuve una mejora de rendimiento cuando contaba con 310 capas pues obtuve una eficiencia del 98.14%. No obstante, siento que es importante resaltar que al momento de probar nuevamente números que ya había colocado me percate que esta va cambiando y puedo decir que con este valor de capas el porcentaje de eficiencia se mantuvo y únicamente aumentó en un 0.01
2. La profundidad del algoritmo. Agreguen una capa escondida más al algoritmo. Este es un ejercicio extremadamente importante!
  - a. ¿Cómo cambia la precisión de validación?
    - i. Al momento de incrementar las capas pude observar que el modelo disminuyo un poco su eficiencia pues ahora contaba con una eficiencia del 97.70% .
  - b. ¿Qué hay del tiempo que se tarda en ejecutar? Pista: deben tener cuidado con las formas de los pesos y los sesgos.
    - i. El tiempo de ejecución aumenta pues ahora cada época para ser entrenada se lleva 5 segundos, pues ahora esta parte nos dio 25s.
3. El ancho y la profundidad del algoritmo. Agregue cuantas capas sean necesarias para llegar a 5 capas escondidas. Es más, ajuste el ancho del algoritmo conforme lo encuentre más conveniente.
  - a. ¿Cómo cambia la precisión de validación?
    - i. La precisión del modelo disminuyó nuevamente pues ahora cuenta con una precisión del 97.52%. Pero me pude percatar de que al momento de irle aplicando más y más capas el modelo iba reduciendo su eficiencia pues llegué a probar hasta con 14 capas y estas me dieron una eficiencia del 97.34%.

- b. ¿Qué hay del tiempo de ejecución?
  - i. En este caso el tiempo de ejecución por época era más tardado pues llegó a tomar 37s en completar el entrenamiento. Al seguir aumentando las capas nos podemos dar cuenta que el tiempo de entrenamiento aumenta haciendo que el tiempo de ejecución sea mayor. Esto lo pude corroborar al probar con 14 capas escondidas
- 4. Experimenten con las funciones de activación. Intenten aplicar una transformación sigmooidal a ambas capas. La activación sigmooidal se obtiene escribiendo "sigmoid".
  - a. Al momento de probar esta activación manteniendo el número de capas de 310 y 5 épocas, me pude percatar que la precisión del modelo redujo significativamente pues esta me dio 97.04% número al cual no llegué ni probando con las 14 capas de activación ReLu. Por otra parte el tiempo de ejecución por épocas es más tardada pues esta con solo 2 capas escondidas entrenar las épocas nos se tomó 25 segundos, resultado que habíamos obtenido con 5 capas ReLu.  
En base a lo anterior podemos ver que este tipo de activación no es la mejor para este modelo.
- 5. Continúen experimentando con las funciones de activación. Intenten aplicar un ReLu a la primera capa escondida y tanh a la segunda. La activación tanh se obtiene escribiendo "tanh".
  - a. Al realizar estos cambios y mezclar dos tipos de activación diferentes me pude percatar rápidamente que al momento de realizar el entrenamiento este se demora más tiempo por época pues estas se tardaron 28 s en entrenarse lo cual es un poco más a las dos capas "sigmoid" pero sigue siendo superior por mucho a contar con las dos capas "ReLu". Ahora bien al hablar de la precisión está aumentó a un 97.88% siendo superior a todas las muestras realizadas con anterioridad.
- 6. Ajusten el tamaño de la tanda. Prueben con un tamaño de tanda de 10,000.
  - a. ¿Cómo cambia el tiempo requerido?
    - i. Para el entrenamiento del modelo el tiempo se vió significativamente reducido pues pasó de un tiempo de 16s a 6.8s. Sin embargo, no se puede decir que siempre al aumentar la cantidad de tandas habrá un aumento en el rendimiento del modelo, pues esto depende directamente de la capacidad de procesamiento que tengamos.
  - b. ¿Cómo cambia la precisión?
    - i. La precisión se vió afectada pues decreció de 97.91 a 91.98. Esto se puede deber a distintos factores, pero principalmente a que el aumento de tandas hace que el backpropagation pierda estabilidad numérica y sea menos preciso al calcular los gradientes y pesos.
- 7. Ajusten el tamaño de la tanda a 1. Eso corresponde al SGD.
  - a. ¿Cómo cambian el tiempo y la precisión?
    - i. El tiempo de ejecución incrementó significativamente, pues al cambiar el tamaño de tandas a 1 el entrenamiento del modelo tomó aproximadamente 8 minutos, a diferencia de los 15s de las 100

tandas. En tanto a la precisión, este disminuyó un 1.28%, pues anteriormente era de 97.81% y ahora es de 96.63%

- b. ¿Es el resultado coherente con la teoría?
  - i. Sí, el resultado es coherente pues al tener una menor cantidad de tandas las actualizaciones de pesos se basan en gradientes calculados a partir de un solo ejemplo de entrenamiento, lo cual dificulta la convergencia hacia una solución óptima.
8. Ajusten la tasa de aprendizaje. Prueben con un valor de 0.0001.
  - a. ¿Hace alguna diferencia?
    - i. En tanto a la precisión del algoritmo no hubo un cambio significativo pues al cambiar la tasa de aprendizaje el porcentaje fue de 96.44%.
9. Ajusten la tasa de aprendizaje a 0.02.
  - a. ¿Hay alguna diferencia?
    - i. Sí, el porcentaje de precisión disminuyó considerablemente pues de un 97.81% disminuyó a 85.66% por lo que se observa que se deben realizar pruebas con distintos valores de tasa de aprendizaje para mejorar el valor de precisión del modelo.
10. Combine todos los métodos indicados arriba e intenten llegar a una precisión de validación de 98.5% o más.
  - a. Algunos de los métodos que nos ayudaron a elevar el porcentaje de precisión fue la implementación de capas escondidas con funciones de activación y la utilización de 100 tandas. Esta cantidad de tandas fue la óptima después de realizar una serie de pruebas, en donde se logró mejorar no solo la precisión sino también el tiempo de entrenamiento del modelo.

Sin embargo, no fue posible llegar a un 98.5% pero se logró acercar bastante el valor a un 97.9% lo cual indica que el modelo aprendió los patrones de los datos y garantiza que este está generalizado para nuevos datos.