Comprender El Deep Learning Requiere Repensar La Generalización

Juan Gabriel Alejandro
Martin Ugarte

Introducción

⇒ Las redes neuronales profundas suelen tener más parámetros de entrenamiento que la cantidad de muestras con las que entrenan los modelos. ¿Qué es lo que diferencia a un modelo que generaliza bien a otro que no?

⇒ Siguiendo la sabiduría del machine learning, se desea buscar la complejidad correcta del modelo acorde a la información que se tiene a mano. No se quiere elegir un modelo que sea tan simple que no pueda ajustar los datos (underfitting), pero tampoco no se quiere crear un modelo tan complicado ya que se aprenderá los datos de entrenamiento de memoria y no logrará generalizar nuevas entradas (overfitting)

Objetivo

⇒ En este trabajo presentamos un marco experimental simple para definir y comprender la noción de la capacidad efectiva de los modelos de aprendizaje automático, propuesto por Chiyuan Zhang en el paper *Understanding Deep Learning Requires Rethiking Generalization*

Experimentos

⇒ **Setup** Se trabaja con el dataset CIFAR 10 (Krizhevksy Hinton, 2009) donde cada imagen es 32x32 con 3 canales de colores. En el preprocesamiento se dividen los valores de los pixeles por 255 para escalarlos de 0 a 1, se los recorta desde el centro para obtener entradas de 28x28 y se normaliza cada imagen independientemente. Para los modelos se utilizó el Gradiente Estocástico con learning rate 0.01 y momentum 0.9

⇒ **Tests de Randomización** Se randomizan etiquetas respecto del dataset original. Por ejemplo se generan etiquetas aleatorias con probabilidad p, así como también se alteran los pixeles de las imagenes con permutaciones constantes o aleatorias, así como también se generar nuevos a partir de una distribución gausseana con determinada media y varianza.

⇒ **Regularización** Se utilizan regularizadores como Data Augmentation, agregando imagenes adicionales transformando las originales de manera que sean verosímiles, y Weight Decay, que reduce el valor de los parámetros agregando un término de penalización a la función de coste. Son usados para combatir el overfitting para lograr una mayor generalización.

Resultados

⇒ Los regularizadores sugieren una mejora en la performance de la generalización, aunque estos no son la razón fundamental para ella, ya que las redes neuronales continuan trabajando bien cuando estos se remueven.

⇒ Se puede lograr un 0 error en training con la randomización de etiquetas, aunque claramente el accuracy en los sets de testeo no será muy bueno ya que la correlación entre los labels de entrenamiento y de prueba no será válida.

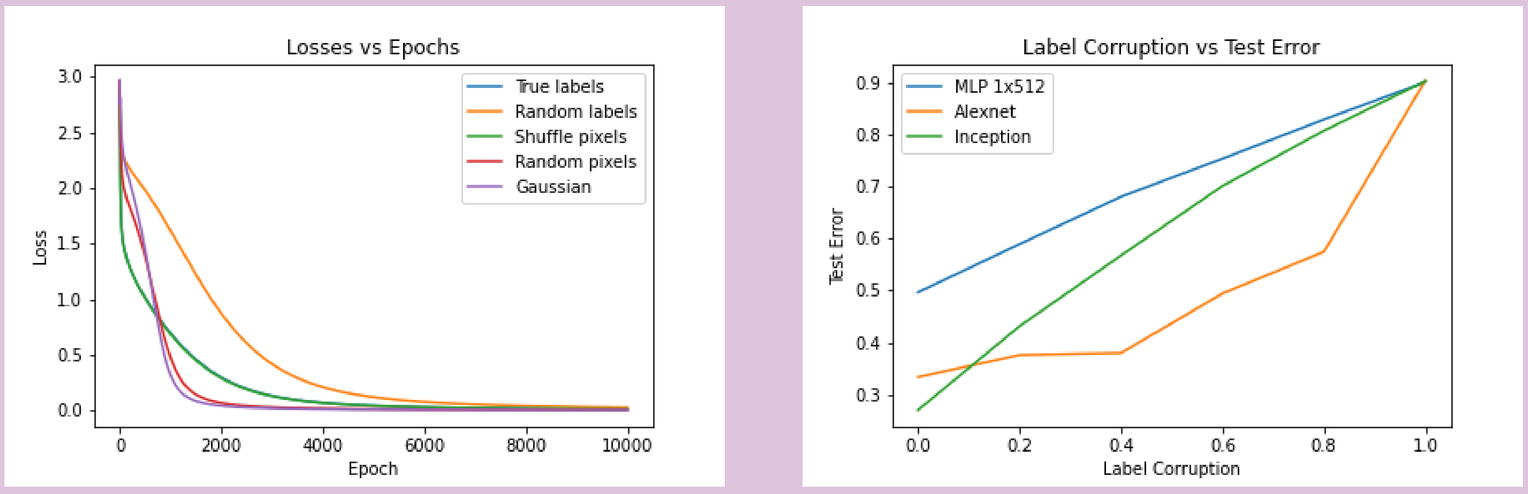

Figura 1: A la izquierda se observa los losses en función de los epochs en cada experimento utilizando la arquitectura MLP. A la derecha se grafica el error de testeo en función de la probabilidad de corrupción de labels para todas las arquitecturas

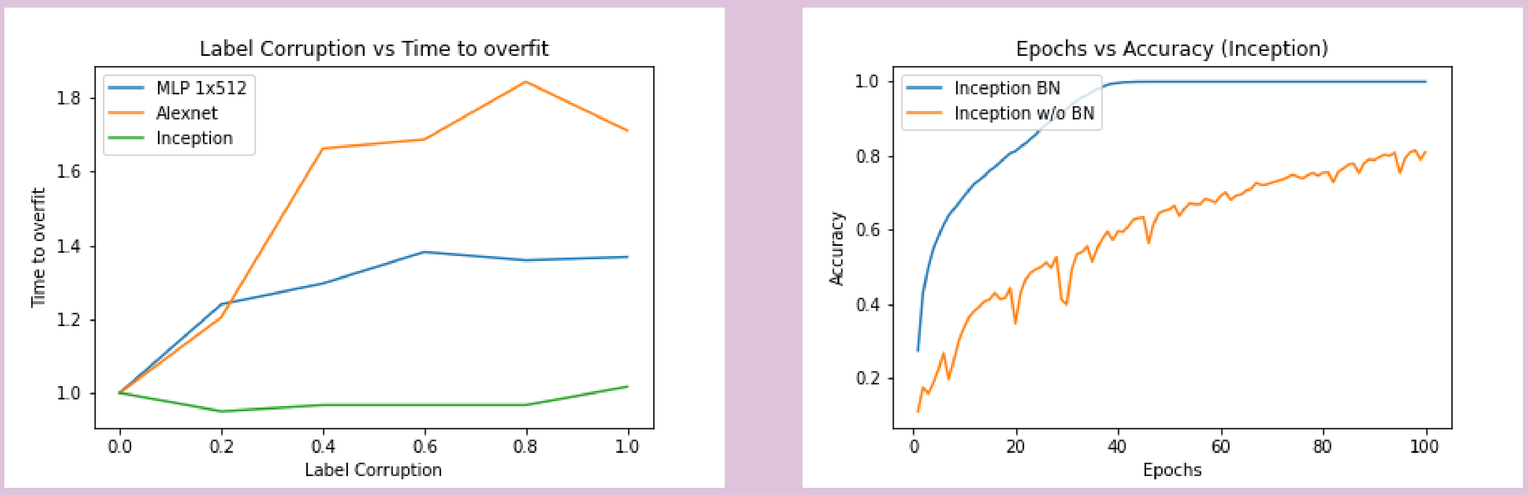

Figura 2: A la izquierda se observa el tiempo que tarda en overfitear cada modeo en función de la probabilidad de corrupción de labels. A la derecha se grafica el accuracy obtenido por cada epoch en el entrenamiento utilizando la arquitectura de Inception

Capacidad efectiva de las redes neuronales

Arquitectura	Train s/ Reg	Test s/ Reg	Train c/ Reg	Test c/ Reg	Train Rand. Labels	Test Rand. Labels
Inception	100.0	73.4	98.25	72.65	100.0	10.0
Inception s/ BatchNorm	86.9	80.6	44.4	44.6	10.3	11.0
Alexnet	81.9	69.6	74.2	67.43	17.3	11.2
MLP 3x512	100.0	48.8	100.0	50.0	88.1	10.2
MLP 1x512	99.0	50.4	35.4	35.1	98.7	10.5

Conclusión

⇒ En el paper se presentan experimentos para definir y entender la noción de capacidad efectiva de los modelos de machine learning. En consecuencia, estos modelos donde se alteran los datos de entrenamiento son lo suficientemente ricos para lograr una buena generalización. Las razones para una fácil optimización deben ser diferentes de la verdadera causa de la generalización.

⇒ La teoría de aprendizaje estadístico lucha para explicar la capacidad de generalización de las redes neuronales. Una medida formal precisa bajo la cual estos enormes modelos son simples aún está por decubrirse.