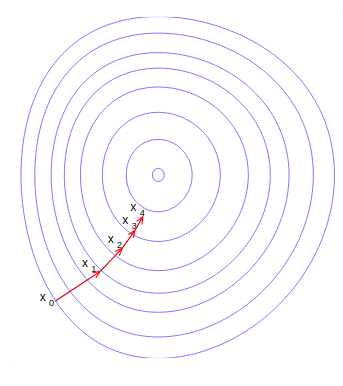
1

**Descenso de Gradientes**

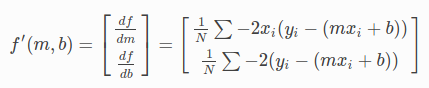
Este es un algoritmo clave que se localiza normalmente en el campo del Aprendizaje Automático, este Descenso de gradientes se localizan en la mayoría de sistemas de inteligencia artificial que se desarrollan hoy en día, la función de coste sera la que nos diga cual es el error para cada una de las combinaciones de nuestros parámetros, ya que matemáticamente esto es una función, entonces su función logarítmica sería encontrar la mayor pendiente en la posición que nos encontramos actualmente, nos dirigimos hacia la pendiente de mayor intensidad, y una vez llegado, repetimos el proceso, esta es el razonamiento del algoritmo del descenso del gradiente, como se vería en el siguiente gráfico.



*Figura 1. imagen bidimensional de la función de gradientes*

***Función del coste***

La pendiente de esta curva nos dice cómo actualizar nuestros parámetros para hacer que el modelo sea más preciso. Calculamos las derivadas parciales de la función de costo con respecto a cada parámetro y almacenamos los resultados en un gradiente, entonces el gradiente se calcularía de la siguiente manera.

 *Figura 2. Esta función indica la pendiente de nuestra función de costo de la posición actual*

Como dijo el blogger Estado Unidense medium,Yudkowsky(2020) "De lejos, el mayor peligro de la inteligencia artificial es que la gente concluye demasiado pronto que la entiende". Entrenar redes neuronales significa que estamos minimizando esta función (llamada función de pérdida). El valor de esta función de pérdida nos da una percepción de cuál es la distancia entre el rendimiento de nuestra red y la perfección en un conjunto de datos determinado.

***Descenso de Gradientes Estocástico***

El descenso de gradientes puede ser lento para ejecutar en conjuntos de datos muy grandes. Debido a que una iteración del algoritmo de descenso de gradiente requiere una predicción para cada instancia en el conjunto de datos de entrenamiento, puede llevar mucho tiempo cuando tiene muchos millones de instancias.

En situaciones en las que tiene grandes cantidades de datos, puede utilizar una variación del descenso de gradiente denominada descenso de gradiente estocástico.

En esta variación, se ejecuta el procedimiento de descenso de gradiente descrito anteriormente, pero la actualización de los coeficientes se realiza para cada instancia de entrenamiento, en lugar de al final del lote de instancias.

El primer paso del procedimiento requiere que el orden del conjunto de datos de entrenamiento sea aleatorio. Esto es para mezclar el orden en que se realizan las actualizaciones de los coeficientes. Debido a que los coeficientes se actualizan después de cada instancia de entrenamiento, las actualizaciones serán ruidosas saltando por todas partes, al igual que la función de costo correspondiente. Al mezclar el orden de las actualizaciones de los coeficientes, aprovecha esta caminata aleatoria y evita que se distraiga o se atasque. El aprendizaje puede ser mucho más rápido con el descenso de gradiente estocástico para conjuntos de datos de entrenamiento muy grandes y, a menudo, solo necesita una pequeña cantidad de pases a través del conjunto de datos para alcanzar un conjunto de coeficientes bueno o suficientemente bueno.

***Conclusión***

La optimización es una gran parte del aprendizaje automático para una buena obtención de datos, ya que en el descenso de gradientes se realiza un proceso de optimización simple que puede usar con muchos algoritmos de aprendizaje automático.

**Cita Textual**

Eleizer Yudkowsky(2020), <https://medium.com/quantyca/gradient-descent-in-deep-learning-b1077b89af81>

**Referencias**

W.A. Gardner, Características de aprendizaje de los algoritmos de descenso de gradiente estocástico: un estudio general, análisis y crítica en el procesamiento de señales, volumen 6, número 2, 1984.

Marcin Andrychowicz y Misha Denil y Sergio Gomez y Matthew W. Hoffman y David Pfau y Tom Schaul y Brendan Shillingford y Nando de Freitas, Aprendiendo a aprender por descenso de gradiente por descenso de gradiente en 1606.04474 arXiv, 2016 .

**Material Electrónico**

Vladimir Meza Ruiz, (2016) Descenso por gradientes en https://turing.iimas.unam.mx/~ivanvladimir/posts/gradient\_descent/