

Họ và tên: Nguyễn Thị Cẩm Lai

MSSV: 20120128

Giải thích ý nghĩa các biến trong file *config*

1. Alpha

- Alpha: được gọi là tốc độ học (learning rate), nó quyết định tốc độ cập nhật các tham số trong quá trình huấn luyện mô hình, điều khiển quá trình hội tụ của Gradient Descent giúp việc học được tối ưu hơn.
- Việc lựa chọn giá trị alpha hợp lý là rất quan trọng để đạt được mô hình tối ưu. Trong quá trình huấn luyện mô hình, ta cần điều chỉnh giá trị alpha sao cho mô hình hội tụ nhanh nhất và không bị vướng phải các vấn đề như overfitting hoặc underfitting.
- Việc lựa chọn giá trị alpha phụ thuộc nhiều vào từng bài toán và phải làm một vài thí nghiệm để chọn ra giá trị tốt nhất
- Hậu quả khi lựa chọn learning rate bất hợp lý:
 - Learning rate quá thấp: trong thực tế, khi việc tính toán trở nên phức tạp, learning rate quá thấp sẽ ảnh hưởng tới tốc độ của thuật toán rất nhiều, thậm chí không bao giờ tới được đích.
 - Learning rate lớn: thuật toán tiến rất nhanh tới gần đích chỉ sau vài vòng lặp. Tuy nhiên, thuật toán không hội tụ được vì bước nhảy quá lớn, khiến nó cứ quần quanh ở đích.



Sự ảnh hưởng của learning rate đến mô hình

2. Iters

- Trong phương pháp Gradient Descent, iter (iteration) thể hiện số lần lặp lại quá trình cập nhật các tham số của mô hình. Mỗi lần lặp, ta tính toán độ dốc của hàm mất mát tại vị trí hiện tại của các tham số, sau đó sử dụng giá trị độ dốc này để điều chỉnh các tham số bằng một khoảng bằng với alpha.
- Việc lặp lại quá trình cập nhật các tham số nhiều lần được thực hiện cho đến khi chúng ta đạt được một giá trị độ dốc đủ nhỏ hoặc một giá trị của hàm mất mát đủ gần với giá trị tối ưu.

- Số lần lặp này phụ thuộc vào nhiều yếu tố như kích thước và độ phức tạp của dữ liệu, độ chính xác mong muốn của kết quả, giá trị alpha, và cấu trúc của mô hình. Thông thường, ta sẽ dừng lại khi đạt được một số lượng lặp đủ lớn hoặc khi giá trị độ dốc đạt đến một ngưỡng nhất định. Tuy nhiên, việc lựa chọn số lần lặp phù hợp là một vấn đề khó và thường được giải quyết bằng cách thực nghiệm để tìm ra số lượng lặp tốt nhất cho từng trường hợp cụ thể.
- Nếu chọn số lượng vòng lặp không hợp lý trong quá trình Gradient Descent thì có thể xảy ra các vấn đề sau:
 - Số lượng vòng lặp quá ít: giá trị của hàm mất mát sẽ không được giảm đáng kể và sẽ không đạt được giá trị tối ưu. Kết quả sẽ không được chính xác và gần giá trị tối ưu mong muốn.
 - Số lượng vòng lặp lớn: quá trình huấn luyện có thể mất quá nhiều thời gian để đạt được kết quả. Điều này có thể làm gián đoạn quá trình phát triển và triển khai các mô hình.