Họ và tên: Mai Minh Quân

MSSV: 20225661

Nguồn tham khảo: Neural Networks and Deep Learning - Michael Nielsen

1.Giới thiệu L1 Regularization

L1 Regularization là một kỹ thuật giúp giảm hiện tượng overfitting trong mô hình học máy bằng cách điều chỉnh hàm chi phí (cost function). Kỹ thuật này liên quan đến việc thêm vào một đại lượng mới vào hàm chi phí, đó là tổng của giá trị tuyệt đối của các trọng số (weights) trong mô hình.

Biểu thức hàm chi phí trong trường hợp này là:

Trong đó,

* ​ là hàm chi phí ban đầu,
* là các trọng số của mô hình và
* là tham số regularization.

Để hiểu rõ hơn cách thức hoạt động của L1 regularization, chúng ta xét đạo hàm của hàm chi phí với đối tượng là các trọng số . Sau khi đạo hàm, ta có biểu thức:

Trong đó, là hàm dấu của trọng số , tức là nếu dương và nếu âm.

Nhờ vào biểu thức này, chúng ta có thể điều chỉnh thuật toán gradient descent để áp dụng L1 regularization. Quy tắc cập nhật trọng số trong L1 regularization sẽ là:

Trong đó, là tốc độ học và ​​ có thể được ước tính thông qua trung bình của mini-batch.

Đối với bias, ta sẽ cập nhật theo công thức:

2. So sánh giữa L1 và L2 Regularization

Cả L1 và L2 regularization đều có tác dụng làm giảm các trọng số của mô hình, giúp giảm sự phức tạp của mô hình và từ đó giảm overfitting. Tuy nhiên, cách thức mà các trọng số giảm đi giữa L1 và L2 có sự khác biệt.

* Trong L1 regularization, so với cách cập nhật trọng số trong thuật toán Gradient Decend các trọng số được thay đổi một lượng cố định () theo hướng về 0. Điều này có nghĩa là mỗi trọng số bị thay đổi một giá trị không phụ thuộc vào giá trị ban đầu của nó, giúp đẩy các trọng số nhỏ về 0.
* Trong khi đó, L2 regularization làm giảm các trọng số theo tỷ lệ trực tiếp với giá trị của chúng. Nghĩa là, trọng số càng lớn thì bị giảm mạnh hơn, còn trọng số nhỏ bị giảm ít hơn.