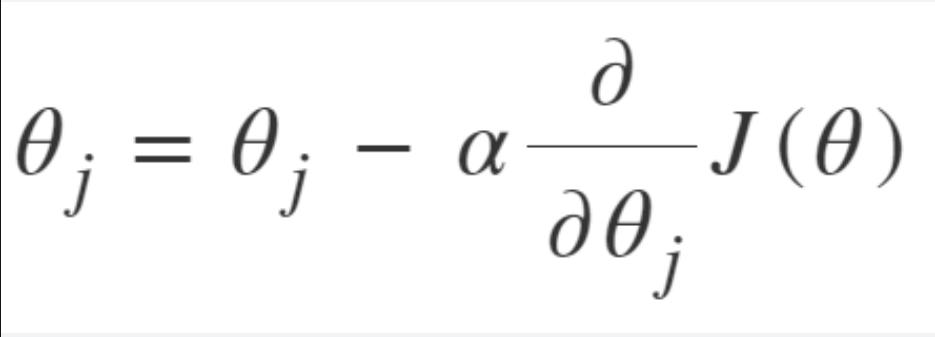
1. Tìm hiểu, so sánh các phương pháp Optimizer trong huấn luyện mô hình học máy;

* Types of gradient descent:
  + Batch Gradient Descent
  + Stochastic Gradient Descent
  + Mini-Batch Gradient Descent

* Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) là một biến thể của Gradient Descent trong quá trình tối ưu hóa tham số trong mô hình máy học. Trong SGD, thay vì cập nhật trọng số (weights) một lần sau mỗi epoch, chúng ta cập nhật trọng số N lần trong mỗi epoch, với N là số lượng điểm dữ liệu trong tập huấn luyện.

Mặt khác, SGD có thể làm giảm tốc độ của một epoch do cần thực hiện nhiều cập nhật trọng số trong một epoch. Tuy nhiên, từ một góc độ khác, SGD có khả năng hội tụ nhanh chóng chỉ sau vài epoch. Công thức của SGD tương tự như Gradient Descent, nhưng được áp dụng trên từng điểm dữ liệu.



1. Tìm hiểu về Continual Learning và Test Production khi xây dựng một giải pháp học máy để giải quyết một bài toán nào đó.