**Bài 1 (3 điểm): làm riêng từng người**

Trình bày một bài nghiên cứu, đánh giá của em về các vấn đề sau:

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

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

1.1 Các phương pháp Optimizer trong huấn luyện mô hình học máy là các thuật toán được sử dụng để tối ưu hóa hàm mất mát và điều chỉnh các trọng số của mô hình.

|  |  |
| --- | --- |
| **Các phương pháp Optimizer** | **Mô tả** |
| **Stochastic Gradient Descent (SGD)** | Dùng một mẫu dữ liệu (mini-batch) ngẫu nhiên để tính gradient và cập nhật trọng số. |
| **Mini-batch Gradient Descent** | Kết hợp giữa hiệu suất của SGD và sự ổn định của Gradient Descent theo lô. |
| **Adaptive Moment Estimation (Adam)** | Kết hợp cả momentum và RMSprop, tự động điều chỉnh learning rate cho từng tham số. |
| **Root Mean Square Propagation (RMSprop)** | Sử dụng trung bình bình phương của gradient để điều chỉnh learning rate. |
| **Adagrad** | Cập nhật learning rate cho từng tham số theo tần suất xuất hiện của nó. |

Đối với mỗi phương pháp, việc lựa chọn phụ thuộc vào bài toán cụ thể và đặc tính của dữ liệu. Ngoài ra, quá trình điều chỉnh siêu tham số như learning rate cũng đóng vai trò quan trọng trong hiệu suất của Optimizer.

1.2 So sánh các phương pháp Optimizer trong huấn luyện mô hình học máy:

|  |  |  |
| --- | --- | --- |
| **Các phương pháp Optimizer** | **Ưu điểm** | **Nhược điểm** |
| **Stochastic Gradient Descent (SGD)** | Tính đơn giản, tốc độ học nhanh trên các tập dữ liệu lớn. | Dễ rơi vào cực tiểu cục bộ, đòi hỏi điều chỉnh learning rate. |
| **Mini-batch Gradient Descent** | Kết hợp hiệu suất và sự ổn định, phổ biến trong thực tế. | Vẫn có thể mắc kẹt ở cực tiểu cục bộ, đòi hỏi tuning kích thước mini-batch. |
| **Adaptive Moment Estimation (Adam)** | Hiệu suất cao, tự động điều chỉnh learning rate. | Có thể quá tinh chỉnh trên một số dữ liệu. |
| **Root Mean Square Propagation (RMSprop)** | Hiệu suất tốt, phù hợp cho dữ liệu không đồng nhất. | Yêu cầu kiểm soát learning rate. |
| **Adagrad** | Phù hợp cho tham số thưa thớt. | Có thể dẫn đến giảm learning rate quá mức. |

2) Continual Learning và Test Production khi xây dựng giải pháp học máy:

**Continual Learning:**

Đặc điểm: Tập trung vào khả năng học liên tục từ dữ liệu mới mà không quên kiến thức cũ.

Thách thức: Catastrophic forgetting - mô hình quên thông tin cũ khi học thông tin mới.

Giải pháp: Elastic Weight Consolidation (EWC), Progressive Neural Networks (PNN).

**Test Production:**

Đặc điểm: Quá trình xây dựng và triển khai giải pháp học máy vào môi trường thực tế để đánh giá hiệu suất.

Quy trình: Chọn lựa dữ liệu, xây dựng mô hình, tinh chỉnh siêu tham số, và kiểm thử trên dữ liệu mới.

Mục tiêu: Đảm bảo mô hình hoạt động hiệu quả và đáp ứng tốt trong môi trường thực tế.

Những sự khác biệt này giúp làm nổi bật tính đặc thù và ưu nhược điểm của từng phương pháp Optimizer trong huấn luyện, cũng như mô tả cách Continual Learning và Test Production đóng vai trò quan trọng trong quá trình xây dựng và triển khai giải pháp học máy.