**Học suốt đời cho các hệ thống học tập nhân tạo**

1. **Introduction**

Các hệ thống máy tính hoạt động trong thế giới thực, tiếp xúc với các luồng thông tin liên tục, do đó được yêu cầu học và nhớ nhiều tác vụ từ các bản phân phối dữ liệu động.

Vấn đề chính của các mô hình tính toán liên quan đến việc học tập suốt đời là chúng dễ bị can thiệp nghiêm trọng, nghĩa là đào tạo một mô hình với thông tin mới can thiếp với những kiến thức cũ đã học trước đây. Hiện tượng này thường dẫn đến giảm hiệu suốt đột ngọt hoặc trong trường hợp xấu nhất là kiến thức cũ bị ghi đè hoàn toàn bởi kiến thức mới

Một sơ đồ học tập giả định rằng tất cả các mẫu đều có sẵn trong giai đoạn đào tạo và do đó, yêu cầu đào tạo lại các tham số mạng trên toàn bộ tập dữ liệu để thích ứng với những thay đổi trong phân phối dữ liệu. Khi được train về sequence task, performance của các mô hình mạng nơ ron thông thường giảm đáng kể đối với các task đã học trước đó khi các tasks mới được học. Mặc dù việc train lại từ đầu về mặt thực dụng có thể giải quyết được tình trạng lãng quên nghiêm trọng (catastrophic forgetting) nhưng phương pháp này rất kém hiệu quả và cản trở việc học các dữ liệu mới trong thời gian thực.

Để khắc phục tình trạng catastrophic forgetting các hệ thống học tập phải cho thấy khả năng thu nhận kiến thức mới và tinh chỉnh kiến thức hiện có trên cơ sở đầu vào liên tục, mặt khác ngăn chặn đầu vào mới có thể can thiệp đáng kể vào kiến thức hiện có. Mức độ mà một hệ thống phải linh hoạt để tích hợp thông tin mới và ổn định theo thứ tự không can thiệp một cách đáng kể ào kiến thức đã được tổng hợp gọi là tình thế tiến thoái lưỡng nan về tính ổn định. => nhiều phương pháp tiếp cận tính toán đã được đề xuất.

Học suốt đời với não người => Bộ não phải thực hiện 2 nhiệm vụ: Khái quát hóa các trải nghiệm và lưu giữ những ký ức cụ thể về các sự kiện giống như từng đợt. Sự thích nghi trong thời gian ngắn cho phép học hỏi nhanh chóng thông tin mới, sau đó sẽ được chuyển giao và tích hợp vào hệ thống neocortical để lưu trữ lâu dài.

Các phương pháp tiếp cận mạng noron thường được thiết kế để dần dần thích ứng với các mẫu dữ liệu tổng hợp được thu thập trong các môi trường được kiểm soát

1. **Các khía cạnh sinh học của học tập suốt đời (biological aspects of lifelong learning)**

**2.1: Độ ổn định - dẻo**

**2.2: Độ dẻo và tính ổn định của Hebbian**

**2.3: Các hệ thống học tập bổ sung**

Não phải có một cơ chế để đồng thời tổng quát hóa các trải nghiệm trong khi vẫn lưu giữ các ký ức theo từng giai đoạn.

**2.4: Học mà không quên**

**3. Học tập suốt đời và sự lãng quên trong neural network**

**3.1: Học tập suốt đời trong Machine Learning**

Học tập suốt đời là một thách thức lâu dài đối với học máy và hệ thống mạng nơ ron. Điều là là do xu hướng các mô hình quên một cách thảm hại kiến thức hiện có khi học từ những kiến thức mới lạ, khiến thông tin mới ghi đè kiến thức đã học trước đó. Ảnh hưởng của catastrophic forgetting đã được nghiên cứu rộng rãi đặc biệt là trong các mạng được học bằng cách sử dụng back-propagation (lan truyền ngược) và mạng Hopfield.

Những nỗ lực ban đầu để giảm thiểu catastrophic forgetting thường bao gồm các hệ thống bộ nhớ lưu trữ dữ liệu đó và thường xuyên phát lại các mẫu cũ xen kẽ các mẫu mới, những phương pháp này vẫn được sử dụng cho đến ngày nay. Tuy nhiên nhược điểm của các hệ thống dựa trên bộ nhớ là chúng yêu cầu lưu trữ thông tin cũ một cách rõ ràng dẫn. Trong kịch bản học tập suốt đời, số lượng nhiệm vụ và mẫu mỗi nhiệm vụ không thể được biết trước nên việc xác định trước một lượng tài nguyên thần kinh đủ để ngăn chặn catastrophic forgetting cần có 3 khía cạnh xác định:

+ Phân bổ tài nguyên neural cho kiến thức mới

+ Sử dụng các đại diện không chồng chéo (non – overlapping) nếu các nguồn lực là cố định

+ Đan xen kiến thức cũ khi thông tin mới được trình bày