**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM  
KHOA ĐÀO TẠO CHẤT LƯỢNG CAO  
NGÀNH CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO TIỂU LUẬN CHUYÊN NGÀNH HTTT**

**ĐỀ TÀI: HỌC SÂU BÁN GIÁM SÁT VÀ ỨNG DỤNG**

**GVHD : TS.NGUYỄN THIÊN BẢO**

**SVTH : LÊ HỮU HẠNH 15110041**

**NGUYỄN ĐỨC QUAN 15110105**

**Tp Hồ Chí Minh, ngày 23 tháng 03 năm 2019**

**-------------------------------------------------------**

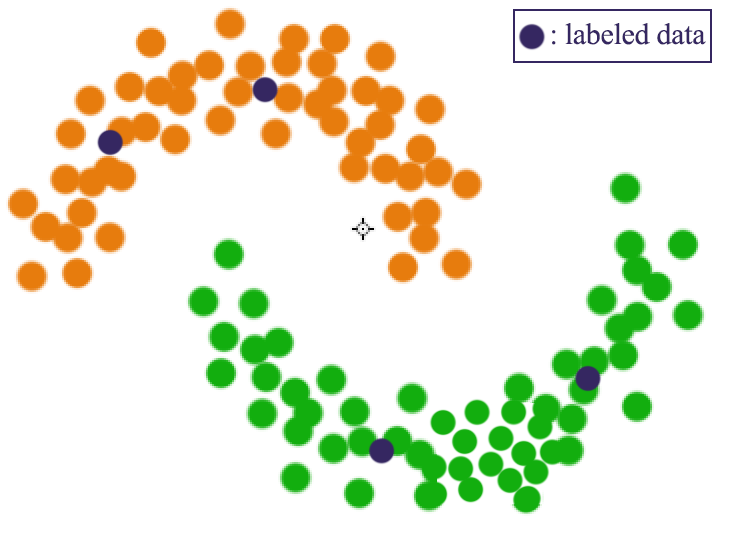
# **CHƯƠNG 1 : GIỚI THIỆU THÀNH PHẦN**

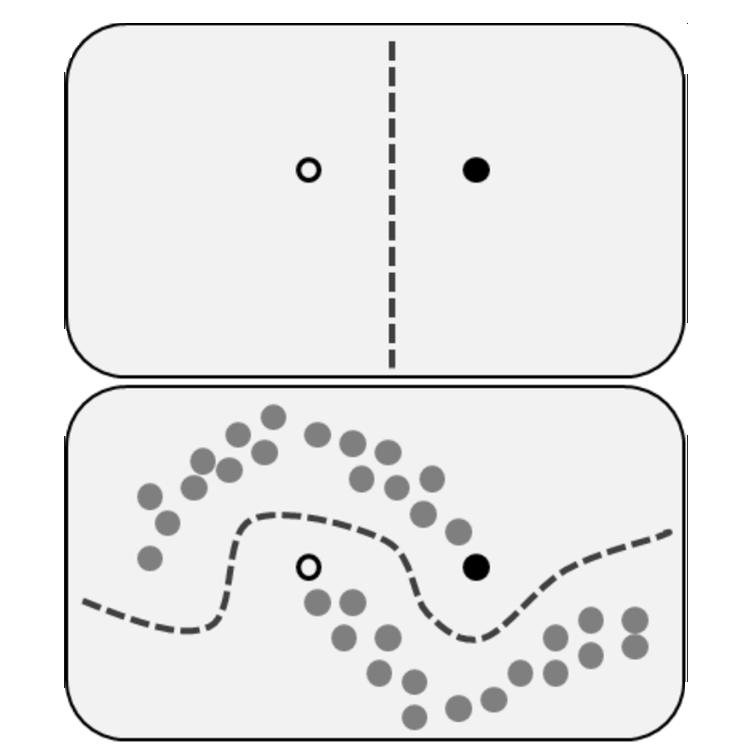
## **1. Bán giám sát**

Học bán giám sát là một lớp các nhiệm vụ và kỹ thuật [học máy](https://en.wikipedia.org/wiki/Machine_learning) cũng sử dụng [dữ liệu](https://en.wikipedia.org/wiki/Data) chưa được gắn nhãn để đào tạo - thường là một lượng nhỏ [dữ liệu được dán nhãn](https://en.wikipedia.org/wiki/Labeled_data) với một lượng lớn dữ liệu chưa được gắn nhãn. Học bán giám sát rơi vào giữa [học tập không giám sát](https://en.wikipedia.org/wiki/Unsupervised_learning) (không có bất kỳ dữ liệu đào tạo được dán nhãn nào) và [học có giám sát](https://en.wikipedia.org/wiki/Supervised_learning) (với dữ liệu đào tạo được dán nhãn hoàn toàn). Nhiều nhà nghiên cứu về máy học đã phát hiện ra rằng dữ liệu không được gắn nhãn, khi được sử dụng cùng với một lượng nhỏ dữ liệu được dán nhãn, có thể tạo ra sự cải thiện đáng kể về độ chính xác của việc học đối với việc học không được giám sát (trong đó không có dữ liệu được dán nhãn), nhưng không có thời gian và chi phí cần thiết cho việc giám sát học tập (nơi tất cả các dữ liệu được dán nhãn).(3)

Việc thu thập dữ liệu được dán nhãn cho một vấn đề học tập thường đòi hỏi một tác nhân người có kỹ năng (ví dụ để phiên âm một đoạn âm thanh) hoặc một thí nghiệm vật lý (ví dụ: xác định cấu trúc 3D của protein hoặc xác định xem có dầu ở một vị trí cụ thể không) . Do đó, chi phí liên quan đến quá trình ghi nhãn có thể khiến tập huấn luyện được dán nhãn đầy đủ là không khả thi, trong khi việc thu thập dữ liệu chưa được gắn nhãn là tương đối rẻ. Trong những tình huống như vậy, học bán giám sát có thể có giá trị thực tiễn lớn. Học bán giám sát cũng là mối quan tâm lý thuyết trong học máy và là một mô hình cho việc học của con người.(3)

Theo trực giác, chúng ta có thể nghĩ về vấn đề học tập như một bài kiểm tra và dữ liệu được dán nhãn là một vài vấn đề mẫu mà giáo viên đã giải quyết trong lớp. Giáo viên cũng cung cấp một tập hợp các vấn đề chưa được giải quyết. Trong môi trường chuyển tải, những vấn đề chưa được giải quyết này là một bài kiểm tra tại nhà và đặc biệt bạn muốn làm tốt với chúng. Trong cài đặt quy nạp, đây là những vấn đề thực hành thuộc loại bạn sẽ gặp trong bài kiểm tra trên lớp.(3)

Hình.1. Học máy và học bán giám sát (Ảnh mạng)

****Hình.2. Một ví dụ về ảnh hưởng của dữ liệu không được gắn nhãn trong học sâu bán giám sát **(**Ảnh mạng**)**

## **2. Học sâu**

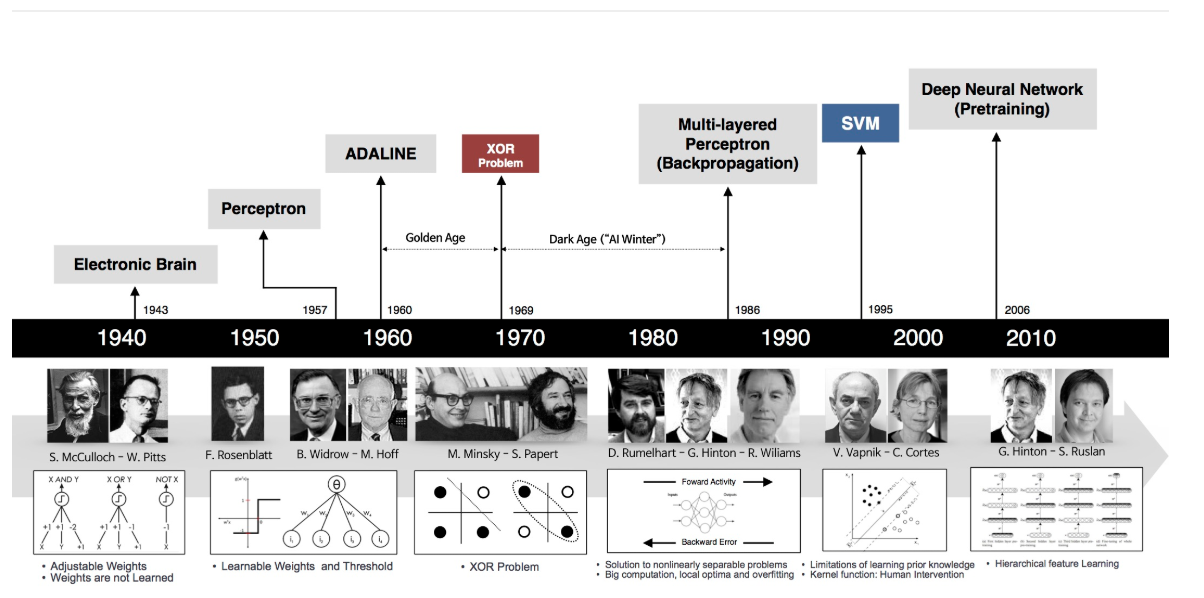
Học sâu ([tiếng Anh](https://vi.wikipedia.org/wiki/Ti%E1%BA%BFng_Anh): *deep learning*) là một chi của ngành [máy học](https://vi.wikipedia.org/wiki/M%C3%A1y_h%E1%BB%8Dc) dựa trên một tập hợp các thuật toán để cố gắng mô hình dữ liệu trừu tượng hóa ở mức cao bằng cách sử dụng nhiều lớp xử lý với cấu trúc phức tạp, hoặc bằng cách khác bao gồm nhiều biến đổi phi tuyến.(3)

Học sâu là một phần của một họ các phương pháp [học máy](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_m%C3%A1y) rộng hơn dựa trên [đại diện học](https://vi.wikipedia.org/w/index.php?title=Learning_representation&action=edit&redlink=1) của dữ liệu. Một quan sát (ví dụ như, một hình ảnh) có thể được biểu diễn bằng nhiều cách như một [vector](https://vi.wikipedia.org/wiki/Kh%C3%B4ng_gian_vect%C6%A1) của các giá trị cường độ cho mỗi điểm ảnh, hoặc một cách trừu tượng hơn như là một tập hợp các cạnh, các khu vực hình dạng cụ thể, vv. Một vài đại diện làm khiến việc học các nhiệm vụ dễ dàng hơn (ví dụ, nhận dạng khuôn mặt hoặc biểu hiện cảm xúc trên khuôn mặt) từ các ví dụ. Một trong những hứa hẹn của học sâu là thay thế các tính năng thủ công bằng các thuật toán hiệu quả đối với [học không có giám sát](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_kh%C3%B4ng_c%C3%B3_gi%C3%A1m_s%C3%A1t)hoặc [nửa giám sát](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_n%E1%BB%ADa_gi%C3%A1m_s%C3%A1t) và tính năng phân cấp.(3)

Các nghiên cứu trong lĩnh vực này cố gắng thực hiện các đại diện tốt hơn và tạo ra các mô hình để tìm hiểu các đại diện này từ dữ liệu không dán nhãn quy mô lớn. Một số đại diện được lấy cảm hứng bởi những tiến bộ trong [khoa học thần kinh](https://vi.wikipedia.org/wiki/Khoa_h%E1%BB%8Dc_th%E1%BA%A7n_kinh) và được dựa trên các giải thích của mô hình xử lý và truyền thông thông tin trong một [hệ thống thần kinh](https://vi.wikipedia.org/wiki/H%E1%BB%87_th%E1%BA%A7n_kinh), chẳng hạn như [mã hóa thần kinh](https://vi.wikipedia.org/w/index.php?title=M%C3%A3_h%C3%B3a_th%E1%BA%A7n_kinh&action=edit&redlink=1) để cố gắng để xác định các mối quan hệ giữa các kích thích khác nhau và các phản ứng liên quan đến thần kinh trong [não](https://vi.wikipedia.org/wiki/N%C3%A3o).(3)

Nhiều kiến trúc học sâu khác nhau như [mạng neuron sâu](https://vi.wikipedia.org/w/index.php?title=M%E1%BA%A1ng_n%C6%A1-ron_s%C3%A2u&action=edit&redlink=1), [mã mạng neuron tích chập sâu](https://vi.wikipedia.org/w/index.php?title=Convolutional_neuron_network&action=edit&redlink=1), [mạng niềm tin sâu](https://vi.wikipedia.org/w/index.php?title=Deep_belief_network&action=edit&redlink=1) và [mạng neuron tái phát](https://vi.wikipedia.org/w/index.php?title=Recurrent_neuron_network&action=edit&redlink=1) đã được áp dụng cho các lĩnh vực như [thị giác máy tính](https://vi.wikipedia.org/wiki/Th%E1%BB%8B_gi%C3%A1c_m%C3%A1y_t%C3%ADnh), [tự động nhận dạng giọng nói](https://vi.wikipedia.org/wiki/Nh%E1%BA%ADn_d%E1%BA%A1ng_ti%E1%BA%BFng_n%C3%B3i), [xử lý ngôn ngữ tự nhiên](https://vi.wikipedia.org/wiki/X%E1%BB%AD_l%C3%BD_ng%C3%B4n_ng%E1%BB%AF_t%E1%BB%B1_nhi%C3%AAn), nhận dạng âm thanh ngôn ngữ và [tin sinh học](https://vi.wikipedia.org/wiki/Tin_sinh_h%E1%BB%8Dc), chúng đã được chứng minh là tạo ra các kết quả rất tốt đối với nhiều nhiệm vụ khác nhau.(3)

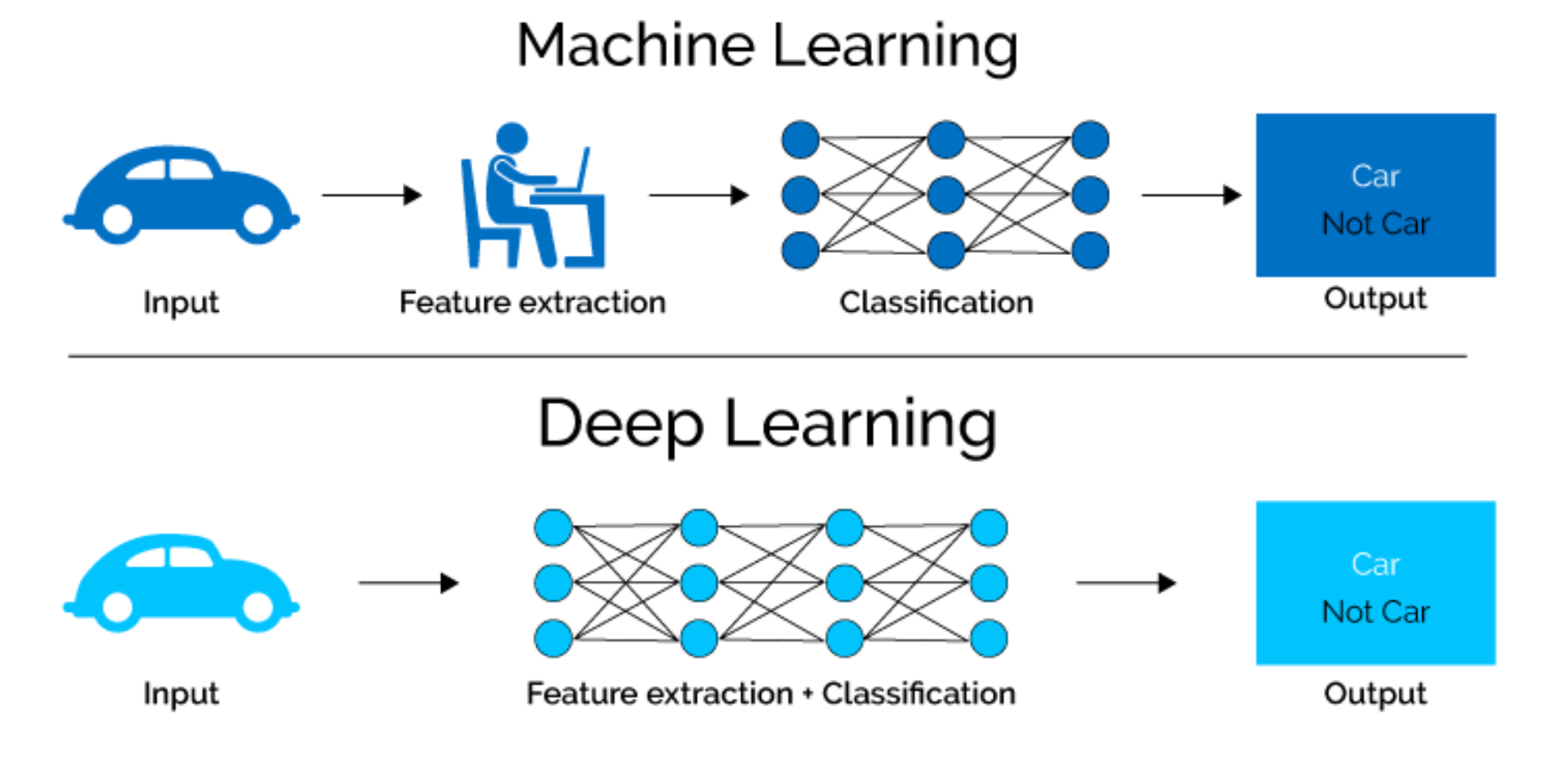
Ngoài ra, *học sâu*đã trở thành một từ ngữ thời thượng, hay một thương hiệu của [mạng neuron](https://vi.wikipedia.org/wiki/M%E1%BA%A1ng_n%C6%A1-ron).

Hình.3. Lịch sử của học sâu(Ảnh mạng)

Có một số cách để mô tả học sâu. Học sâu là một lớp của các [thuật toán](https://vi.wikipedia.org/wiki/Thu%E1%BA%ADt_to%C3%A1n) [máy học](https://vi.wikipedia.org/wiki/M%C3%A1y_h%E1%BB%8Dc) mà:

* Sử dụng một tầng (cascade) nhiều lớp các đơn vị [xử lý phi tuyến](https://vi.wikipedia.org/w/index.php?title=Nonlinear_filter&action=edit&redlink=1) để [trích tách đặc điểm](https://vi.wikipedia.org/w/index.php?title=Feature_extraction&action=edit&redlink=1) và chuyển đổi. Mỗi lớp kế tiếp dùng đầu ra từ lớp trước làm đầu vào. Các thuật toán này có thể được [giám sát](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_c%C3%B3_gi%C3%A1m_s%C3%A1t) hoặc [không cần giám sát](https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_kh%C3%B4ng_c%C3%B3_gi%C3%A1m_s%C3%A1t) và các ứng dụng bao gồm các mô hình phân tích (không có giám sát) và phân loại (giám sát).(3)
* Dựa trên học (không có giám sát) của nhiều cấp các đặc điểm hoặc đại diện của dữ liệu. Các tính năng cao cấp bắt nguồn từ các tính năng thấp cấp hơn để tạo thành một đại diện thứ bậc.(3)
* Là một phần của lĩnh vực máy học rộng lớn hơn về việc học đại diện dữ liệu.(3)
* Học nhiều cấp độ đại diện tương ứng với các mức độ trừu tượng khác nhau; các mức độ hình thành một hệ thống phân cấp của các khái niệm.(3)

Các định nghĩa này có điểm chung là (1) nhiều lớp các đơn vị xử lý phi tuyến và (2) học có giám sát hoặc không có giám sát của biểu diễn đặc tính ở mỗi lớp, với các lớp hình thành một hệ thống các tính năng phân cấp từ thấp đến cao cấp. Các thành phần của một lớp của đơn vị xử lý phi tuyến sử dụng một thuật toán học sâu tùy theo vấn đề cần được giải quyết. Các lớp được sử dụng trong học sâu bao gồm các lớp ẩn của một mạng [neuron nhân tạo](https://vi.wikipedia.org/wiki/M%E1%BA%A1ng_n%C6%A1-ron_nh%C3%A2n_t%E1%BA%A1o) và tập các [công thức mệnh đề](https://vi.wikipedia.org/w/index.php?title=Propositional_formula&action=edit&redlink=1) phức tạp. Chúng cũng có thể bao gồm các biến tiềm ẩn được tổ chức thành các lớp chọn lọc trong các mô hình thể sinh (có khả năng sinh ra) sâu(3)

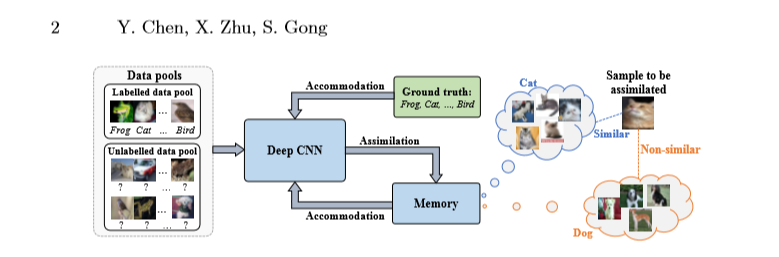
Hình.4. Quy trình của học sâu(Ảnh mạng)

# **CHƯƠNG 2 : HỌC SÂU BÁN GIÁM SÁT**

## **1. Giới thiệu**

Học tập bán giám sát (Semi-supervised learning) nhằm mục đích tăng hiệu suất mô hình bằng cách sử dụng số lượng lớn dữ liệu không được gắn nhãn khi chỉ có một số lượng dữ liệu có nhãn hạn chế [4,37]. Đó là động lực mà dữ liệu không có nhãn hiệu có sẵn ở quy mô lớn nhưng dữ liệu được dán nhãn là khan hiếm do chi phí ghi nhãn cao. Lược đồ học tập này hữu ích và hữu ích đối với nhiều ứng dụng như tìm kiếm hình ảnh [6], phân loại trang web [2], truy xuất tài liệu [21], bộ gen [29], v.v. Trong tài liệu SSL, thuật toán SSL đơn giản nhất là tự đào tạo nơi mô hình đích được đào tạo từng bước bằng dữ liệu tự dán nhãn bổ sung được đưa ra bởi các dự đoán riêng của mô hình với con số cao [21,2,25]. Phương pháp này có khuynh hướng tuyên truyền lỗi trong học mô hình do dự đoán sai về sự kết hợp cao. Các phương pháp phổ biến khác bao gồm Transductive SVM [10,3] và phương pháp graphbased [39,1], tuy nhiên, có khả năng chịu thiệt hại từ khả năng mở rộng kém đến dữ liệu không được đánh dấu quy mô lớn do tối ưu hóa không hiệu quả.(4)

Gần đây, các phương pháp SSL dựa trên mạng nơron [23,35,15,12,30,24,19,26,16, 9,32] bắt đầu thống trị tiến trình nhờ khả năng học tập biểu diễn mạnh mẽ của mạng thần kinh sâu. Hầu hết các phương pháp này thường sử dụng mạng đào tạo cập nhật để xây dựng hình phạt bổ sung không giám sát (4)

Hình.5. Illustration of the memory-assisted semi-supervised deep learning framework(4)

để cho phép học tập bán giám sát. Chúng ta xem xét rằng loại lược đồ SSL sâu này không tối ưu với điều kiện khả năng ghi nhớ của các mạng sâu thường không đầy đủ và không bị ngăn cản để thể hiện kiến ​​thức được tích luỹ trong quá trình học lặp lại [34]. Để thúc đẩy kiến ​​thức như vậy, chúng ta giới thiệu một cơ chế bộ nhớ vào quá trình đào tạo mạng sâu để cho phép học tập bán giám sát từ dữ liệu đào tạo có nhãn có kích thước nhỏ và dữ liệu đào tạo không có nhãn có kích thước lớn. Theo tinh thần của lý thuyết của Piaget về khả năng học tập liên tục của con người [7], chúng ta hướng tới thiết kế một sơ đồ SSL cho phép mô hình sâu tìm hiểu thêm từ bộ nhớ của nó (đồng hóa) và tự điều chỉnh để tối ưu dữ liệu đào tạo (chỗ ở) theo cách gia tăng. Để đạt được mục đích này, chúng ta xây dựng một khung học tập sâu được giám sát bán bộ nhớ mới hỗ trợ bộ nhớ: Mạng nơ-ron sâu được hỗ trợ bộ nhớ (MA-DNN) như được minh họa trong hình 1. MA-DNN được đặc trưng bởi sự tương tác giữa đồng hóa và chỗ ở giữa mạng và một mô-đun bộ nhớ ngoài.(4)

Chìa khóa để thiết kế khung công tác của chúng ta là hai khía cạnh: (1) biểu diễn tính năng phân biệt đối xử cấp lớp và sự không chắc chắn suy luận mạng được tích lũy dần dần trong một mô-đun bộ nhớ ngoài; (2) thông tin ghi nhớ này được sử dụng để đồng hóa các mẫu hình ảnh mới đến và tạo ra một mất mát bộ nhớ không giám sát thông tin để hướng dẫn việc học mạng cùng với sự mất phân lớp giám sát.(4)

Đóng góp của chúng ta là hai lần: (1) Chúng ta đề xuất khai thác bộ nhớ của học mô hình để cho phép học tập sâu bán giám sát từ một ít dữ liệu đào tạo có nhãn và nhiều dữ liệu đào tạo không có nhãn, đồng thời hoàn toàn áp dụng quy trình đào tạo từ đầu đến cuối. Điều này trái ngược với hầu hết các phương thức SSL sâu hiện có, thường bỏ qua bộ nhớ của việc học mô hình. (2) Chúng ta xây dựng một mạng lưới thần kinh sâu bộ nhớ mới hỗ trợ bộ nhớ (MA-DNN) đặc trưng bởi một cơ chế bộ nhớ. Chúng ta giới thiệu một mất mát bộ nhớ không giám sát tương thích với sự mất mát phân loại theo tiêu chuẩn giám sát để cho phép học tập bán giám sát. Các thí nghiệm so sánh mở rộng thể hiện những ưu điểm của mô hình MA-DNN được đề xuất của chúng ta qua nhiều phương pháp học sâu rộng được giám sát hiện đại nhất.(4)