Chuyên đề. Khảo mắt mạng noron và ứng dụng.

Trừ tượng

Kể từ khi đề xuất một thuật toán học nhanh cho các mạng niềm tin sâu sắc vào năm 2006, các kỹ thuật học sâu

đã thu hút sự quan tâm nghiên cứu ngày càng tăng vì khả năng vốn có của họ là khắc phục nhược điểm của

các thuật toán truyền thống phụ thuộc vào các tính năng được thiết kế bằng tay. Phương pháp học tập sâu cũng đã được tìm thấy

phù hợp để phân tích dữ liệu lớn với các ứng dụng thành công cho thị giác máy tính, nhận dạng mẫu, nhận dạng giọng nói,

xử lý ngôn ngữ tự nhiên, và các hệ thống khuyến nghị. Trong bài báo này, chúng tôi thảo luận về một số học tập sâu được sử dụng rộng rãi

kiến trúc và ứng dụng thực tế của họ. Tổng quan cập nhật được cung cấp trên bốn kiến ​​trúc học tập sâu,

cụ thể là bộ mã hóa tự động, mạng nơ ron tích chập, mạng lưới niềm tin sâu sắc và máy Boltzmann bị hạn chế. Khác nhau

các loại mạng lưới thần kinh sâu được khảo sát và những tiến bộ gần đây được tóm tắt. Ứng dụng học sâu

kỹ thuật trên một số khu vực được chọn (nhận dạng giọng nói, nhận dạng mẫu và thị giác máy tính) được nêu bật. Một

danh sách các chủ đề nghiên cứu trong tương lai cuối cùng được đưa ra với các biện minh rõ ràng.

Điều khoản chỉ mục

Autoencoder, Mạng thần kinh chuyển đổi, Học sâu, Mạng niềm tin sâu, Boltzmann bị hạn chế

Máy móc

I. GIỚI THIỆU

Kỹ thuật học máy đã được áp dụng rộng rãi trong nhiều lĩnh vực như nhận dạng mẫu, tự nhiên

xử lý ngôn ngữ và học tập tính toán. Với kỹ thuật học máy, máy tính được trang bị

khả năng hành động mà không được lập trình rõ ràng, xây dựng các thuật toán có thể học hỏi từ dữ liệu,

và đưa ra quyết định hoặc dự đoán dựa trên dữ liệu. Trong những thập kỷ qua, học máy đã mang lại rất nhiều

ảnh hưởng đến cuộc sống hàng ngày của chúng tôi với các ví dụ bao gồm tìm kiếm web hiệu quả, hệ thống tự lái, thị giác máy tính,

và nhận dạng ký tự quang học. Ngoài ra, bằng cách áp dụng các phương pháp học máy, nhân tạo ở cấp độ con người

trí thông minh (AI) cũng đã được cải thiện, xem [101], [137], [165] để biết thêm các cuộc thảo luận. Tuy nhiên, khi nó

nói đến các cơ chế xử lý thông tin của con người (ví dụ: lời nói và tầm nhìn), hiệu suất của truyền thống

kỹ thuật máy học là xa thỏa đáng. Lấy cảm hứng từ cấu trúc phân cấp sâu sắc của lời nói của con người

hệ thống nhận thức và sản xuất, khái niệm thuật toán học sâu đã được giới thiệu vào cuối thế kỷ 20.

Những đột phá về học tập sâu đã đạt được từ năm 2006 khi Hinton đề xuất một cuốn tiểu thuyết có cấu trúc sâu

kiến trúc học tập được gọi là mạng lưới niềm tin sâu sắc (DBN) [59]. Thập kỷ qua đã chứng kiến ​​sự phát triển nhanh chóng của các kỹ thuật học sâu với những tác động đáng kể đến việc xử lý tín hiệu và thông tin. Nghiên cứu về thần kinh

các hệ thống cũng hỗ trợ phát triển các mô hình mạng sâu [75]. Trái ngược với học máy truyền thống

và phương pháp tiếp cận trí tuệ nhân tạo, các công nghệ học tập sâu gần đây đã được phát triển ồ ạt với

ứng dụng thành công để nhận dạng giọng nói, xử lý ngôn ngữ tự nhiên (NLP), truy xuất thông tin, tính toán

tầm nhìn và phân tích hình ảnh

Khái niệm học sâu bắt nguồn từ nghiên cứu trên các mạng thần kinh nhân tạo (ANNs) [60]. ANN có

trở thành một lĩnh vực nghiên cứu tích cực trong vài thập kỷ qua [63], [162], [166], [167], [175]. Để xây dựng một tiêu chuẩn

mạng lưới thần kinh (NN), điều cần thiết là sử dụng tế bào thần kinh để tạo ra các kích hoạt có giá trị thực và, bằng cách điều chỉnh

trọng lượng, các NN hành xử như mong đợi. Tuy nhiên, tùy thuộc vào các vấn đề, quá trình đào tạo NN có thể mất

chuỗi nhân quả dài của các giai đoạn tính toán. Backpropagation là một thuật toán giảm độ dốc hiệu quả, có

đã đóng một vai trò quan trọng trong NN từ năm 1980. Nó đào tạo ANN với phương pháp học tập có sự giám sát của giáo viên.

Mặc dù độ chính xác đào tạo cao, hiệu suất của truyền ngược khi áp dụng cho dữ liệu thử nghiệm có thể

không thỏa đáng Vì backpropagation dựa trên thông tin độ dốc cục bộ với điểm ban đầu ngẫu nhiên,

thuật toán thường bị mắc kẹt trong tối ưu cục bộ. Hơn nữa, nếu kích thước của dữ liệu đào tạo không đủ lớn, NN

sẽ phải đối mặt với vấn đề quá mức. Do đó, các thuật toán học máy hiệu quả khác như hỗ trợ

máy vectơ (SVM), tăng cường và hàng xóm gần nhất (KNN) đã được áp dụng để đạt được tối ưu toàn cầu

với mức tiêu thụ điện năng thấp hơn. Năm 2006, Hinton [59] đã đề xuất một phương pháp đào tạo mới (được gọi là lớp-khôn ngoan-greedylearning) đánh dấu sự ra đời của các kỹ thuật học sâu. Ý tưởng cơ bản của việc học tập theo tầng lớp

là học tập không giám sát nên được thực hiện cho đào tạo trước mạng trước từng lớp tiếp theo

đào tạo. Bằng cách trích xuất các tính năng từ các đầu vào, kích thước dữ liệu được giảm và biểu diễn nhỏ gọn là

do đó thu được. Sau đó, xuất các tính năng sang lớp tiếp theo, tất cả các mẫu sẽ được dán nhãn và mạng

sẽ được tinh chỉnh với dữ liệu được dán nhãn. Lý do cho sự phổ biến của học tập sâu là gấp đôi: một mặt,

phát triển các kỹ thuật phân tích dữ liệu lớn chỉ ra rằng vấn đề quá mức trong dữ liệu đào tạo có thể là một phần

giải quyết; mặt khác, quy trình đào tạo trước khi học không giám sát sẽ chỉ định ban đầu không ngẫu nhiên

các giá trị cho mạng. Do đó, tối thiểu địa phương tốt hơn có thể đạt được sau quá trình đào tạo và nhanh hơn

tốc độ hội tụ có thể đạt được.

Cho đến nay, nghiên cứu về các kỹ thuật học sâu đã khuấy động rất nhiều sự chú ý và một loạt thú vị

kết quả đã được báo cáo trong các tài liệu. Kể từ năm 2009, cuộc thi ImageNet, đã thu hút rất nhiều người

nhóm nghiên cứu thị giác máy tính trên toàn thế giới từ cả học viện và ngành công nghiệp. Năm 2012, nghiên cứu

nhóm được dẫn dắt bởi Hinton đã giành chiến thắng trong cuộc thi Phân loại hình ảnh ImageNet bằng cách sử dụng các phương pháp học tập sâu

[86]. Lần đầu tiên nhóm Hinton tham dự cuộc thi và kết quả của họ tốt hơn 10% so với

nơi thứ hai. Cả Google và Yahoo đã cập nhật các công cụ tìm kiếm hình ảnh của họ dựa trên việc học sâu

kiến trúc với những cải tiến lớn trong tìm kiếm chính xác. Baidu cũng thành lập Viện nghiên cứu sâu (IDL)

vào năm 2013 và mời Andrew Ng, phó giáo sư tại Đại học Stanford, làm Nhà khoa học trưởng. Tháng Ba

Năm 2016, một trận đấu trò chơi cờ vây đã được tổ chức tại Hàn Quốc bởi dự án học sâu của Google (được gọi là DeepMind) giữa

người chơi AI của họ AlphaGo và một trong những người chơi mạnh nhất thế giới, Lee Se-dol [140]. Hóa ra AlphaGo,

áp dụng các kỹ thuật học sâu, cho thấy sức mạnh đáng ngạc nhiên và đánh bại Lee Se-dol với tỷ lệ 4: 1. Ngoài ra, học sâu

Các thuật toán cũng cho thấy hiệu suất vượt trội trong việc dự đoán hoạt động của các phân tử thuốc tiềm năng và

ảnh hưởng của đột biến DNA không mã hóa đến biểu hiện gen.

Với sự phát triển nhanh chóng của các kỹ thuật tính toán, một khung mạnh mẽ đã được ANN cung cấp với

kiến trúc sâu cho học tập có giám sát. Nói chung, thuật toán học sâu bao gồm một hệ thống phân cấp

kiến trúc với nhiều lớp, mỗi lớp tạo thành một đơn vị xử lý thông tin phi tuyến tính. Trong bài báo này, chúng tôi

chỉ thảo luận về kiến trúc sâu trong NN. Mạng nơ ron sâu (DNN), sử dụng kiến trúc sâu trong NN, có thể biểu diễn các hàm có độ phức tạp cao hơn nếu số lượng lớp và đơn vị trong một lớp tăng lên.

Được cung cấp đủ bộ dữ liệu đào tạo và mô hình phù hợp, phương pháp học sâu có thể giúp con người thiết lập

chức năng ánh xạ để thuận tiện cho hoạt động. Trong bài báo này, bốn kiến trúc sâu chính được thu hồi và khác

các phương thức (ví dụ: mã hóa thưa thớt) cũng được thảo luận ngắn gọn. Ngoài ra, một số tiến bộ gần đây trong lĩnh vực sâu

học tập được mô tả. Mục đích của bài viết này là cung cấp một đánh giá và giới thiệu kịp thời về các công nghệ học tập sâu và

ứng dụng của họ. Nó nhằm mục đích cung cấp cho độc giả một nền tảng về các kiến trúc học tập sâu khác nhau

và cũng là sự phát triển mới nhất cũng như những thành tựu trong lĩnh vực này. Phần còn lại của bài báo được tổ chức như sau.

Trong Phần II - V, bốn kiến trúc học sâu chính, đó là các máy Boltzmann (RBM) bị hạn chế, sâu

Các mạng niềm tin (DBN), bộ mã hóa tự động (AE) và mạng nơ ron tích chập (CNN), được xem xét, tương ứng.

So sánh được thực hiện trong số các kiến trúc sâu sắc và những phát triển gần đây về các thuật toán này sẽ được thảo luận.

Các ứng dụng của những kiến trúc sâu này được nêu bật trong Phần VI. Kết luận và chủ đề nghiên cứu trong tương lai

được trình bày trong Phần VII.

II. KIẾN TRÚC HỌC TẬP: MÁY BOLTZMANN GIỚI HẠN

A. Động lực

Trong phần này, một đánh giá ngắn gọn về RBM được đưa ra. RBM được sử dụng rộng rãi trong các mạng học tập sâu trên tài khoản của

tầm quan trọng lịch sử của họ và sự đơn giản tương đối. RBM lần đầu tiên được đề xuất như một khái niệm của Smolensky, và

đã trở nên nổi bật kể từ khi Hinton công bố tác phẩm của mình [59] vào năm 2006. RBM đã được sử dụng để tạo ra stochastic

các mô hình ANN có thể tìm hiểu phân phối xác suất liên quan đến đầu vào của chúng. RBM bao gồm một

biến thể của máy Boltzmann (BMs). BM có thể được hiểu là NN với các đơn vị xử lý ngẫu nhiên được kết nối

hai chiều. Vì rất khó để tìm hiểu các khía cạnh của phân phối xác suất không xác định, RBM đã được đề xuất

để đơn giản hóa cấu trúc liên kết của mạng và để nâng cao hiệu quả của mô hình. Nó cũng được công nhận rằng một

RBM là một loại trường ngẫu nhiên Markov đặc biệt với các đơn vị có thể nhìn thấy ngẫu nhiên trong một lớp và có thể quan sát được ngẫu nhiên

đơn vị trong lớp khác.

B. Cấu trúc và thuật toán

Như được hiển thị trong Hình 1, các nơ-ron bị hạn chế tạo thành đồ thị lưỡng cực trong RBM. Có thể thấy rằng có

là một kết nối đầy đủ giữa các đơn vị hiển thị và các đơn vị ẩn, trong khi không có kết nối nào tồn tại giữa các đơn vị từ

cùng một lớp [165]. Để huấn luyện RBM, bộ lấy mẫu Gibbs được thông qua. Bắt đầu với một trạng thái ngẫu nhiên trong một lớp

và thực hiện lấy mẫu Gibbs, chúng ta có thể tạo dữ liệu từ RBM. Khi các trạng thái của các đơn vị trong một lớp

được đưa ra, tất cả các đơn vị trong các lớp khác sẽ được cập nhật. Quá trình cập nhật này sẽ tiếp tục cho đến khi cân bằng

phân phối đạt được. Tiếp theo, các trọng số trong một RBM có được bằng cách tối đa hóa khả năng của RBM này.

Cụ thể, lấy độ dốc của xác suất đăng nhập của dữ liệu huấn luyện, các trọng số có thể được cập nhật theo

đến:



Trong đó ωij đại diện cho trọng số giữa đơn vị hiển thị i và đơn vị ẩn j. hv

0

tôi h

0

j

tôi và hv∞

tôi h

j

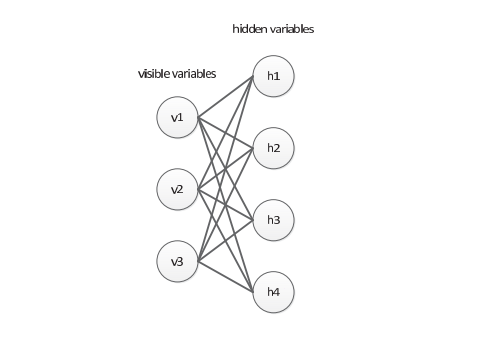
tôi là

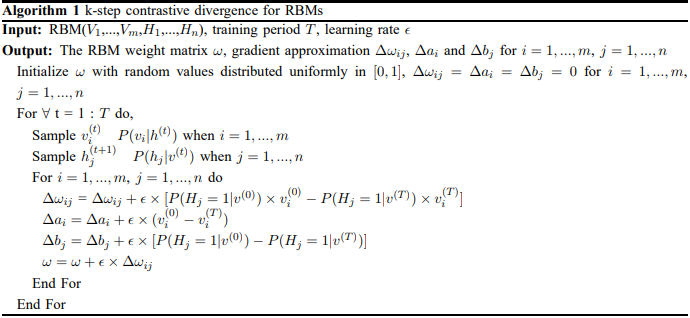
tương quan khi các đơn vị nhìn thấy và ẩn ở lớp thấp nhất và lớp cao nhất, tương ứng. Các

bằng chứng chi tiết có thể được tìm thấy trong [59]. Cần lưu ý rằng quá trình đào tạo sẽ hiệu quả hơn khi sử dụng

thuật toán phân kỳ tương phản dựa trên độ dốc (CD). Thuật toán CD cho đào tạo RBM được phát triển bởi

Hinton năm 2002 [56]. Quy trình của thuật toán CD k-step được đưa ra trong Thuật toán 1.





Giả sử rằng sự khác biệt giữa mô hình và phân phối mục tiêu không lớn, chúng ta có thể sử dụng các mẫu

được tạo bởi chuỗi Gibbs để xấp xỉ độ dốc âm. Lý tưởng nhất là khi chiều dài của chuỗi tăng lên, nó

đóng góp cho khả năng giảm và có xu hướng bằng không [12]. Tuy nhiên, trong [147], chúng ta có thể thấy rằng ước tính

của gradient không thể đại diện cho chính gradient. Hơn nữa, hầu hết các thành phần CD và độ dốc loglikabilities tương ứng có dấu bằng [45]. Do đó, một thuật toán thực tế hơn được gọi là phân kỳ tương phản liên tục

đã được đề xuất trong [115]. Trong phương pháp này, các tác giả đề nghị theo dõi trạng thái của các chuỗi liên tục hơn là

tìm kiếm giá trị ban đầu của chuỗi Gibbs Markov tại một vectơ dữ liệu đã cho. Các trạng thái của ẩn và nhìn thấy

các đơn vị trong chuỗi liên tục được gia hạn sau khi cập nhật từng trọng lượng. Theo cách này, ngay cả một tỷ lệ học tập nhỏ

sẽ không gây ra nhiều khác biệt giữa các bản cập nhật và trạng thái chuỗi liên tục trong khi mang lại chính xác hơn

C. Biến thể của RBM

Ngày nay, RBM đang đóng một vai trò quan trọng trong các ứng dụng khác nhau như mô hình chủ đề, tính chiều

giảm, lọc cộng tác, phân loại và tính năng học tập. Ví dụ, một RBM có thể được sử dụng để mã hóa

dữ liệu và sau đó được áp dụng cho việc học tập không giám sát để hồi quy hoặc phân loại. Ngoài ra, RBM có thể là

được sử dụng như một mô hình thế hệ. Chúng ta có thể tính toán phân phối chung của các đơn vị P (v, h) có thể nhìn thấy và ẩn

Luật Bayes. Xác suất có điều kiện của một đơn vị p (h | v) cũng có thể được tính bằng RBM. Do đó, một

RBM cũng có thể được sử dụng như một mô hình phân biệt đối xử.

Nói chung, RBM được sử dụng làm trình trích xuất tính năng trong quy trình đào tạo trước cho các nhiệm vụ phân loại. Tuy nhiên,

các tính năng được trích xuất bởi RBM trong học tập không giám sát có thể không hữu ích trong quá trình học có giám sát. Trong

Ngoài ra, việc lựa chọn các tham số, rất quan trọng đối với hiệu suất của các thuật toán học tập, cũng sẽ mang lại

Khó khăn. Để xử lý những vấn đề này, máy Boltzmann (DRBM) bị hạn chế phân biệt đối xử đã được đề xuất bởi

Larochelle và Bengio vào năm 2008. Hơn nữa, để học trực tuyến với các bộ dữ liệu lớn, mô hình của DBRM lai

(HDRBM) hoạt động tốt do lợi thế kết hợp của cả học tập chung và phân biệt đối xử. Tuy nhiên, trong các nhiệm vụ phân loại đa nhãn, hiệu suất của RBM không thỏa đáng. Mnih et al. [115] đề xuất

cái gọi là máy Boltzmann bị hạn chế có điều kiện (CRBM) để cải thiện hiệu suất hơn nữa. Trong khi đó,

trong chuỗi thời gian chiều cao, CRBM có thể được sử dụng làm mô hình thế hệ phi tuyến tính. Trong [153], một vô hướng

mô hình được thiết lập với các biến có thể nhìn thấy có giá trị thực và các biến tiềm ẩn nhị phân. Trong mô hình này, các biến có thể nhìn thấy

tại một số bước thời gian cuối cùng có thể bị ảnh hưởng trực tiếp bởi các biến tiềm ẩn và có thể nhìn thấy ở mỗi bước thời gian. Với

tài sản này, suy luận trực tuyến có thể được thực hiện hiệu quả hơn bởi CRBM. Ngoài ra, học hỏi từ thời gian

loạt, CRBM có thể có được các đại diện phân tán phong phú để đảm bảo hiệu quả chính xác

suy luận

Gần đây, một DRBM khép kín (được gọi là FE-RBM) đã được Elfwing phát triển dựa trên một tiểu thuyết phân biệt đối xử

thuật toán học tập [41]. Trong FE-RBM, đầu ra cho bất kỳ vectơ đầu vào và lớp nào được tính theo

năng lượng tự do âm của RBM. Mục tiêu học tập đạt được thông qua việc giảm thiểu việc đào tạo bình phương

lỗi sử dụng một phương pháp giảm độ dốc ngẫu nhiên. Hơn nữa, lấy cảm hứng từ nghiên cứu trước đó, năng lượng tự do

được chia tỷ lệ theo hằng số dựa trên kích thước mạng để cải thiện độ mạnh của xấp xỉ hàm trong

FE-RBM.

Khi RBM được áp dụng cho các khu vực như nhận dạng hình ảnh và giọng nói, hiệu suất của chúng có thể bị suy giảm nghiêm trọng

bởi những tiếng ồn trong dữ liệu [55]. Năm 2012, Tang et al. [152] đã giới thiệu một mô hình hiện đại, Boltzmann mạnh mẽ

máy (RoBM), có thể được sử dụng để đối phó với tiếng ồn và tiếng kêu trong nhận dạng hình ảnh. Với RoBM,

một sự khái quát hóa tốt hơn có thể đạt được bằng cách loại bỏ ảnh hưởng của các pixel bị hỏng. Được đào tạo với nhãn mác

dữ liệu với tiếng ồn sử dụng thuật toán nghiêng không giám sát, mô hình RoBM cũng có thể tìm hiểu cấu trúc không gian của

các chất thải So với các thuật toán truyền thống, các RoBM đã cho thấy hiệu suất nâng cao ở nhiều mức độ khác nhau

các ứng dụng như inpainting hình ảnh và nhận dạng khuôn mặt.

Là một yếu tố chính trong phân phối Boltzmann, lần đầu tiên, nhiệt độ được xem xét trong

mô hình đồ họa của DBN của Li et al. [97]. Các máy Boltzmann (TRBM) bị hạn chế dựa trên nhiệt độ là

đề xuất trong đó nhiệt độ đóng vai trò là một tham số độc lập được điều chỉnh. Phân tích lý thuyết cho thấy rằng

nhiệt độ là yếu tố chính kiểm soát tính chọn lọc của các nơ-ron bắn trong các lớp ẩn. Nó đã được chứng minh

hiệu suất của các TRBM được đề xuất có thể được tăng cường bằng cách cài đặt đúng thông số độ sắc nét của

chức năng hậu cần. Do mức độ linh hoạt bổ sung được giới thiệu, TRBM có thể thu được kết quả chính xác hơn.

Hơn nữa, nghiên cứu cũng cung cấp một số hiểu biết về RBM từ quan điểm vật lý chỉ ra

rằng có thể tồn tại một số mối quan hệ giữa nhiệt độ và một số NN ngoài đời thực.

III. KIẾN TRÚC HỌC TẬP: DEEP MẠNG NIỀM TIN

A. Động lực

Như đã đề cập trong phần trước, các biến ẩn và nhìn thấy không độc lập lẫn nhau [165]. Đến

khám phá sự phụ thuộc giữa các biến này, vào năm 2006, Hinton đã xây dựng các DBN bằng cách xếp chồng một ngân hàng

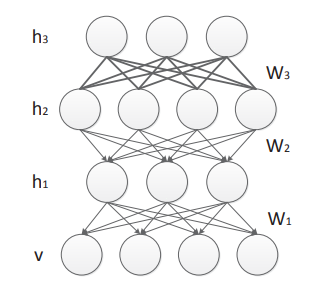
RBM. Cụ thể, các DBN bao gồm nhiều lớp biến ngẫu nhiên và tiềm ẩn và có thể được xem xét

như một hình thức đặc biệt của mô hình phát sinh xác suất Bayes. So với ANN, DBN có hiệu quả hơn,

đặc biệt là khi áp dụng cho các vấn đề với dữ liệu không nhãn.

B. Cấu trúc và thuật toán

Sơ đồ của mô hình được hiển thị bên dưới trong Hình 2



Có thể thấy trong Hình 2 rằng trong một DBN, cứ hai lớp liền kề tạo thành một RBM. Lớp có thể nhìn thấy của mỗi

RBM được kết nối với lớp ẩn của RBM trước đó và hai lớp trên cùng là không định hướng. Chỉ đạo

kết nối giữa lớp trên và lớp dưới theo cách từ trên xuống. Các lớp RBM khác nhau trong một

DBN được đào tạo tuần tự: các RBM thấp hơn được đào tạo trước, sau đó là các RBM cao hơn. Sau khi tính năng được trích xuất

bởi RBM hàng đầu, chúng sẽ được truyền trở lại các lớp thấp hơn [30]. So với một RBM, xếp chồng

mô hình sẽ tăng giới hạn trên của khả năng đăng nhập, hàm ý khả năng học tập mạnh hơn [5].

Quá trình đào tạo của một DBN có thể được chia thành hai giai đoạn: giai đoạn tiền huấn luyện và giai đoạn tinh chỉnh.

Trong giai đoạn tiền đào tạo, đào tạo dựa trên học tập không giám sát được thực hiện theo hướng đi xuống cho tính năng

khai thác; trong khi ở giai đoạn tinh chỉnh, thuật toán lên xuống học tập có giám sát được thực hiện để tiếp tục

điều chỉnh các tham số mạng. Chúng tôi lưu ý rằng hiệu suất được cải thiện của các DBN có thể được quy cho phần lớn

đến giai đoạn tiền đào tạo trong đó các trọng số ban đầu của mạng được học từ cấu trúc của dữ liệu đầu vào.

So với các trọng số được khởi tạo ngẫu nhiên, các trọng số này gần với tối ưu toàn cầu hơn và do đó có thể mang lại

hiệu suất tốt hơn.

Thuật toán CD được giới thiệu trong phần trước có thể được sử dụng để huấn luyện trước một DBN. Hiệu suất, tuy nhiên,

thường không đạt yêu cầu đặc biệt là khi dữ liệu đầu vào được kẹp. Để khắc phục vấn đề này, một thuật toán học tập lớp lớp tham lam đã được giới thiệu nhằm tối ưu hóa các trọng số của một DBN tại tuyến tính phức tạp theo thời gian theo kích thước và độ sâu của mạng [59]. Trong thuật toán học từng lớp tham lam, các RBM tạo thành một

DBN được đào tạo tuần tự. Cụ thể, lớp hiển thị của RBM thấp nhất được đào tạo trước với h

(0) làm đầu vào.

Các giá trị trong lớp hiển thị sau đó được nhập vào các lớp ẩn trong đó xác suất kích hoạt P (h | v) của

các biến ẩn được tính toán. Các đại diện thu được trong RBM trước đó sẽ được sử dụng làm đào tạo

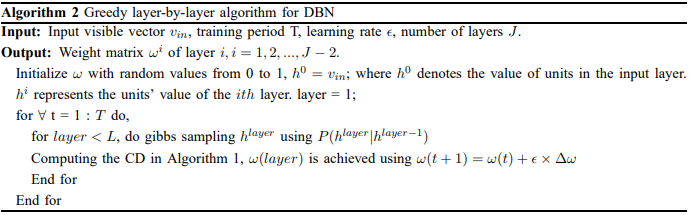
dữ liệu cho RBM tiếp theo và quá trình đào tạo này tiếp tục cho đến khi tất cả các lớp được duyệt qua. Vì trong thuật toán này,

xấp xỉ hàm khả năng chỉ được yêu cầu trong một bước, thời gian đào tạo đã được đáng kể

giảm. Vấn đề thiếu sáng thường xảy ra trong các mạng sâu cũng có thể được khắc phục trong giai đoạn tiền huấn luyện

quá trình. Thuật toán đào tạo trước này còn được gọi là thuật toán đào tạo không giám sát từng lớp một. Dành cho

Rõ ràng, chúng tôi đã cung cấp quy trình thực hiện trong Thuật toán 2



Trong giai đoạn tinh chỉnh, các DBN được đào tạo với dữ liệu được dán nhãn bằng thuật toán lên xuống tương phản

phiên bản của thuật toán đánh thức [57]. Để tìm ra ranh giới danh mục của mạng, một bộ nhãn được đặt

đến lớp trên cùng cho quá trình học tập công nhận trọng lượng. Ngoài ra, thuật toán backpropagation được sử dụng để hoàn thiện các trọng số với dữ liệu được dán nhãn [149]. So với thuật toán đánh thức ban đầu, thuật toán lên xuống

không bị các vấn đề về tính trung bình của chế độ có thể mang lại trọng lượng nhận dạng kém.

Tóm lại, quy trình đào tạo của một DBN bao gồm một quy trình đào tạo trước từng lớp không được giám sát

thực hiện theo cách từ dưới lên và quy trình tinh chỉnh lên xuống có giám sát. Quá trình đào tạo trước có thể là

được coi là tính năng học tập thông qua đó có thể đạt được giá trị ban đầu tốt hơn cho các trọng số và lên xuống

thuật toán sau đó được sử dụng để điều chỉnh toàn bộ mạng. Thật đáng để đề cập rằng với DBN, dữ liệu không được gắn nhãn là

chế biến hiệu quả. Hơn nữa, các vấn đề quá mức và thiếu chất lượng cũng có thể tránh được [30].