Deep Learning là gì? Deep Learning là một tập hợp con của Machine Learning, bản thân nó nằm trong lĩnh vực trí tuệ nhân tạo.

**Sự khác biệt giữa AI, Machine Learning và Deep Learning là gì?  
Trí tuệ nhân tạo là nghiên cứu về cách chế tạo những cỗ máy có khả năng thực hiện các nhiệm vụ thường đòi hỏi trí thông minh của con người.**

AI(Artificial Intelligence) bao gồm nhiều lĩnh vực nghiên cứu, từ thuật toán di truyền đến các hệ thống chuyên gia và cung cấp phạm vi cho các lập luận về những gì cấu thành AI.

Trong lĩnh vực nghiên cứu AI, Machine Learning đã đạt được thành công đáng kể trong những năm gần đây – cho phép máy tính vượt qua hoặc tiến gần đến việc kết hợp hiệu suất của con người trong các lĩnh vực từ nhận dạng khuôn mặt đến nhận dạng giọng nói và ngôn ngữ.

Machine Learning là quá trình dạy máy tính thực hiện một nhiệm vụ, thay vì lập trình nó làm thế nào để thực hiện nhiệm vụ đó từng bước một.

Khi kết thúc đào tạo, một hệ thống Machine Learning sẽ có thể đưa ra dự đoán chính xác khi được cung cấp dữ liệu.

Điều đó nghe có vẻ khô khan, nhưng những dự đoán đó có thể trả lời liệu một miếng trái cây trong ảnh là chuối hay táo, nếu một người đang băng qua trước một chiếc xe tự lái, cho dù việc sử dụng sách từ trong câu liên quan đến bìa mềm hoặc đặt phòng khách sạn, cho dù email là thư rác hay nhận dạng giọng nói đủ chính xác để tạo chú thích cho video YouTube.

Machine Learning thường được chia thành học có giám sát, trong đó máy tính học bằng ví dụ từ dữ liệu được gắn nhãn và học không giám sát, trong đó các máy tính nhóm các dữ liệu tương tự và xác định chính xác sự bất thường.

Deep Learning là một tập hợp con của Machine Learning, có khả năng khác biệt ở một số khía cạnh quan trọng so với Machine Learning nông truyền thống, cho phép máy tính giải quyết một loạt các vấn đề phức tạp không thể giải quyết được.

Một ví dụ về một nhiệm vụ Machine Learning đơn giản, nông cạn có thể dự đoán doanh số bán kem sẽ thay đổi như thế nào dựa trên nhiệt độ ngoài trời. Việc đưa ra dự đoán chỉ sử dụng một vài tính năng dữ liệu theo cách này là tương đối đơn giản và có thể được thực hiện bằng cách sử dụng một kỹ thuật Machine Learning gọi là hồi quy tuyến tính với độ dốc giảm dần.

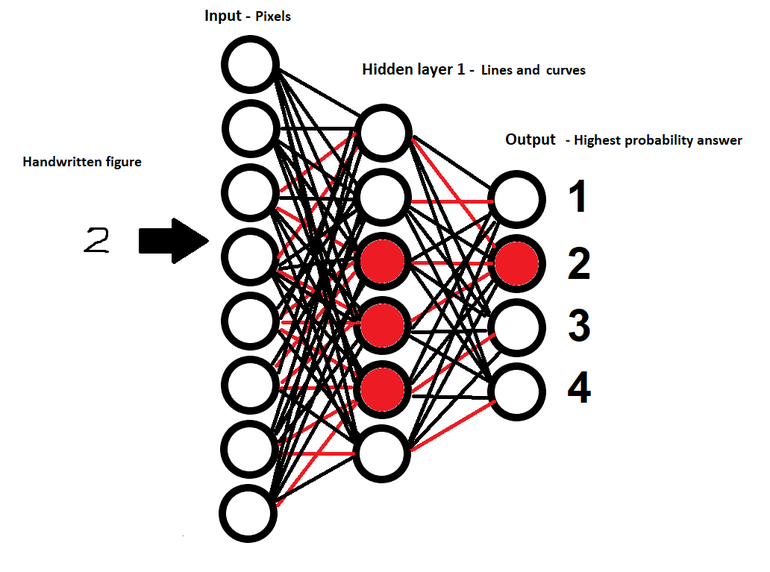
Vấn đề là hàng loạt vấn đề trong thế giới thực không phù hợp với những mô hình đơn giản như vậy. Một ví dụ về một trong những vấn đề thực tế phức tạp này là nhận ra các số viết tay.

Để giải quyết vấn đề này, máy tính cần phải có khả năng đối phó với sự đa dạng lớn trong cách thức trình bày dữ liệu. Mỗi chữ số từ 0 đến 9 có thể được viết theo vô số cách: kích thước và hình dạng chính xác của mỗi chữ số viết tay có thể rất khác nhau tùy thuộc vào người viết và trong hoàn cảnh nào.

Đối phó với sự biến đổi của các tính năng này và sự lộn xộn tương tác lớn hơn giữa chúng, là nơi học tập sâu và mạng lưới thần kinh sâu trở nên hữu ích.

Mạng lưới thần kinh là các mô hình toán học có cấu trúc được lấy cảm hứng lỏng lẻo từ bộ não.

Mỗi nơ-ron trong mạng nơ-ron là một hàm toán học lấy dữ liệu thông qua đầu vào, biến đổi dữ liệu đó thành dạng dễ điều chỉnh hơn và sau đó phun ra thông qua đầu ra. Bạn có thể nghĩ về các nơ-ron trong một mạng lưới thần kinh như được sắp xếp theo lớp, như hình dưới đây.



**Hình ảnh: Nick Heath / ZDNet**

Tất cả các mạng thần kinh đều có một lớp đầu vào, trong đó dữ liệu ban đầu được đưa vào và một lớp đầu ra, tạo ra dự đoán cuối cùng. Nhưng trong một mạng lưới thần kinh sâu, sẽ có nhiều “lớp tế bào” ẩn giữa các lớp đầu vào và đầu ra, mỗi lớp cho dữ liệu vào nhau. Do đó, thuật ngữ “Deep” trong “Deep Learning” và “mạng lưới thần kinh sâu”, nó liên quan đến số lượng lớn các lớp ẩn – thường lớn hơn ba – tại trung tâm của các mạng thần kinh này.

Sơ đồ đơn giản hóa ở trên hy vọng sẽ giúp cung cấp một ý tưởng về cách cấu trúc một mạng lưới thần kinh đơn giản. Trong ví dụ này, mạng đã được đào tạo để nhận ra các số liệu viết tay, chẳng hạn như số 2 được hiển thị ở đây, với lớp đầu vào được cung cấp các giá trị đại diện cho các pixel tạo thành hình ảnh của một chữ số viết tay và lớp đầu ra dự đoán số viết tay nào đã được hiển thị trong hình ảnh.

Trong sơ đồ trên, mỗi vòng tròn đại diện cho một nơ-ron trong mạng, với các nơ-ron được tổ chức thành các lớp thẳng đứng.

Như bạn có thể thấy, mỗi nơ-ron được liên kết với mọi nơ-ron ở lớp sau, thể hiện thực tế là mỗi nơ-ron tạo ra một giá trị vào mỗi nơ-ron ở lớp tiếp theo. Màu sắc của các liên kết trong sơ đồ cũng khác nhau. Các màu khác nhau, đen và đỏ, thể hiện tầm quan trọng của các liên kết giữa các nơ-ron. Các liên kết màu đỏ là những liên kết có ý nghĩa lớn hơn, có nghĩa là chúng sẽ khuếch đại giá trị khi nó đi qua giữa các lớp. Đổi lại, sự khuếch đại giá trị này có thể giúp kích hoạt tế bào thần kinh mà giá trị đang được đưa vào.

Một nơ-ron có thể được cho là đã được kích hoạt khi tổng các giá trị được đưa vào nơ-ron này vượt qua ngưỡng đã đặt. Trong sơ đồ, các tế bào thần kinh được kích hoạt có màu đỏ. Kích hoạt này có nghĩa là khác nhau theo lớp. Trong “Lớp ẩn 1” được hiển thị trong sơ đồ, một nơ ron kích hoạt có thể có nghĩa là hình ảnh của hình viết tay chứa một tổ hợp pixel nhất định giống với đường nằm ngang ở đầu số viết tay 7. Theo cách này, “Lớp ẩn 1 “Có thể phát hiện nhiều đường và đường cong câu chuyện cuối cùng sẽ kết hợp với nhau thành hình viết tay đầy đủ.

Một mạng lưới thần kinh thực tế có thể sẽ có cả hai lớp ẩn và nhiều nơ-ron hơn trong mỗi lớp. Ví dụ: “Lớp ẩn 2” có thể được cung cấp các đường và đường cong nhỏ được xác định bởi “Lớp ẩn 1” và phát hiện cách chúng kết hợp để tạo thành các hình dạng có thể nhận biết, tạo thành các chữ số, như toàn bộ vòng lặp dưới cùng của sáu. Bằng cách cung cấp dữ liệu chuyển tiếp giữa các lớp theo cách này, mỗi lớp ẩn tiếp theo xử lý các tính năng ngày càng cao hơn.

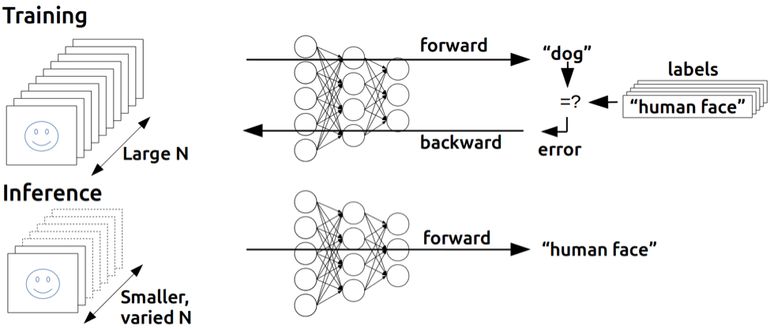
Như đã đề cập, tế bào thần kinh được kích hoạt trong lớp đầu ra của sơ đồ có một ý nghĩa khác. Trong trường hợp này, tế bào thần kinh được kích hoạt tương ứng với số lượng mạng thần kinh ước tính nó được hiển thị trong hình ảnh của một chữ số viết tay mà nó được cung cấp làm đầu vào.

Như bạn có thể thấy, đầu ra của một lớp là đầu vào của lớp tiếp theo trong mạng, với dữ liệu chảy qua mạng từ đầu vào đến đầu ra.

Nhưng làm thế nào để nhiều lớp ẩn này cho phép một máy tính xác định bản chất của một chữ số viết tay? Nhiều lớp tế bào thần kinh này về cơ bản cung cấp một cách để mạng lưới thần kinh xây dựng một hệ thống phân cấp thô gồm các tính năng khác nhau tạo nên chữ số viết tay trong câu hỏi. Chẳng hạn, nếu đầu vào là một mảng các giá trị đại diện cho các pixel riêng lẻ trong hình ảnh của hình viết tay, lớp tiếp theo có thể kết hợp các pixel này thành các đường và hình dạng, lớp tiếp theo kết hợp các hình dạng đó thành các đặc điểm riêng biệt như các vòng lặp trong 8 hoặc tam giác trên trong 4, và như vậy. Bằng cách xây dựng một bức tranh về các tính năng này, các mạng thần kinh hiện đại có thể xác định – với độ chính xác rất cao – con số tương ứng với một chữ số viết tay. Tương tự, các loại mạng thần kinh sâu khác nhau có thể được đào tạo để nhận diện khuôn mặt trong hình ảnh hoặc để phiên âm lời nói bằng văn bản.

Quá trình xây dựng hệ thống phân cấp ngày càng phức tạp này của các tính năng của số viết tay không có gì ngoài các pixel được mạng học. Quá trình học tập được thực hiện bằng cách mạng có thể thay đổi tầm quan trọng của các liên kết giữa các nơ-ron trong mỗi lớp. Mỗi liên kết có một giá trị đính kèm được gọi là trọng số, nó sẽ sửa đổi giá trị được tạo ra bởi một nơron khi nó truyền từ lớp này sang lớp kế tiếp. Bằng cách thay đổi giá trị của các trọng số này và một giá trị liên quan được gọi là sai lệch, có thể nhấn mạnh hoặc làm giảm tầm quan trọng của các liên kết giữa các nơ-ron trong mạng.

Ví dụ, trong trường hợp mô hình nhận dạng chữ số viết tay, các trọng số này có thể được sửa đổi để nhấn mạnh tầm quan trọng của một nhóm pixel cụ thể tạo thành một dòng hoặc một cặp các đường giao nhau tạo thành 7.



**Một minh họa về cấu trúc của một mạng lưới thần kinh và cách đào tạo hoạt động.**

Mô hình học được các liên kết giữa các nơ-ron rất quan trọng trong việc đưa ra dự đoán thành công trong quá trình đào tạo. Ở mỗi bước trong quá trình đào tạo, mạng sẽ sử dụng một hàm toán học để xác định mức độ chính xác của dự đoán mới nhất của nó so với dự kiến. Hàm này tạo ra một loạt các giá trị lỗi, do đó hệ thống có thể sử dụng để tính toán cách mô hình nên cập nhật giá trị của các trọng số được gắn vào mỗi liên kết, với mục đích cuối cùng là cải thiện độ chính xác của các dự đoán của mạng. Mức độ mà các giá trị này sẽ được thay đổi được tính bởi một chức năng tối ưu hóa, chẳng hạn như giảm độ dốc và những thay đổi đó được đẩy lùi trên toàn mạng vào cuối mỗi chu kỳ đào tạo trong một bước gọi là lan truyền ngược.

Trải qua nhiều, rất nhiều chu kỳ đào tạo và với sự trợ giúp của việc điều chỉnh tham số thủ công không thường xuyên, mạng sẽ tiếp tục nue để tạo dự đoán tốt hơn và tốt hơn cho đến khi nó đạt gần với độ chính xác cao nhất. Tại thời điểm này, ví dụ, khi các chữ số viết tay có thể được nhận ra với độ chính xác hơn 95%, mô hình Deep Learning có thể nói là đã được đào tạo.

Về cơ bản, Deep Learning cho phép Machine Learning giải quyết một loạt các vấn đề phức tạp mới – chẳng hạn như nhận dạng hình ảnh, ngôn ngữ và lời nói – bằng cách cho phép máy móc tìm hiểu cách các tính năng trong dữ liệu kết hợp thành các dạng trừu tượng ngày càng cao hơn. Ví dụ: trong nhận dạng khuôn mặt, cách các pixel trong hình ảnh tạo ra các đường và hình dạng, cách các đường và hình dạng đó tạo ra các đặc điểm khuôn mặt và cách các đặc điểm khuôn mặt này được sắp xếp thành một khuôn mặt.

**Tại sao nó được gọi là Deep Learning?**

Như đã đề cập, độ sâu đề cập đến số lượng các lớp ẩn, thường là hơn ba, được sử dụng trong các mạng lưới thần kinh sâu.

**Làm thế nào mà Deep Learning được sử dụng?**

Đối với nhiều nhiệm vụ, để nhận biết và tạo hình ảnh, lời nói và ngôn ngữ và kết hợp với học tăng cường để phù hợp với hiệu suất của con người trong các trò chơi từ cổ đại, như Go, đến hiện đại, như Dota 2 và Quake III.

Hệ thống học tập sâu là một nền tảng của các dịch vụ trực tuyến hiện đại. Các hệ thống như vậy được Amazon sử dụng để hiểu những gì bạn nói – cả lời nói và ngôn ngữ bạn sử dụng – với trợ lý ảo Alexa hoặc Google để dịch văn bản khi bạn truy cập trang web tiếng nước ngoài.

Mỗi tìm kiếm của Google sử dụng nhiều hệ thống Machine Learning, để hiểu ngôn ngữ trong truy vấn của bạn thông qua việc cá nhân hóa kết quả của bạn, vì vậy những người đam mê câu cá tìm kiếm “bass” không bị ngập trong kết quả về guitar.

Nhưng ngoài những biểu hiện rất rõ ràng về máy móc và học tập sâu, các hệ thống như vậy đang bắt đầu tìm thấy một ứng dụng trong mọi ngành công nghiệp. Những ứng dụng này bao gồm: tầm nhìn máy tính cho xe không người lái, máy bay không người lái và robot giao hàng; nhận dạng và tổng hợp ngôn ngữ và ngôn ngữ cho chatbot và robot dịch vụ; nhận dạng khuôn mặt để giám sát ở các nước như Trung Quốc; giúp các bác sĩ X quang chọn ra các khối u trong tia X, giúp các nhà nghiên cứu phát hiện ra các chuỗi di truyền liên quan đến các bệnh và xác định các phân tử có thể dẫn đến các loại thuốc hiệu quả hơn trong chăm sóc sức khỏe; cho phép bảo trì dự đoán về cơ sở hạ tầng bằng cách phân tích dữ liệu cảm biến IoT; củng cố tầm nhìn máy tính giúp siêu thị Amazon Go không thu tiền có thể cung cấp phiên âm và dịch thuật chính xác hợp lý cho các cuộc họp kinh doanh – danh sách này vẫn tiếp tục.

go-gallery-nội thất-v534007255.jpg  
Cửa hàng Amazon Go dựa vào nhận dạng hình ảnh được hỗ trợ bằng cách tìm hiểu sâu để phát hiện những gì người mua hàng mua.



**Khi nào thì bạn nên sử dụng Deep Learning**

Khi dữ liệu của bạn phần lớn không có cấu trúc và bạn có rất nhiều dữ liệu.

Các thuật toán Deep Learning có thể lấy dữ liệu lộn xộn và không có nhãn rộng rãi – chẳng hạn như video, hình ảnh, bản ghi âm thanh và văn bản – và áp đặt đủ thứ tự cho dữ liệu đó để đưa ra dự đoán hữu ích, xây dựng hệ thống phân cấp các tính năng tạo nên con chó hoặc con mèo một hình ảnh hoặc âm thanh tạo thành một từ trong lời nói.

Sự bùng nổ của IoT sẽ thay đổi cách phân tích dữ liệu  
Xin lỗi, AI nói chung vẫn còn rất xa  
10 công nghệ này rất có thể sẽ giúp cứu hành tinh Trái đất  
Cuộc chiến của Google về những kẻ thù sâu sắc: Khi cuộc bầu cử hiện ra, nó chia sẻ hàng tấn video giả mạo AI

**Deep Learning hay giải quyết những vấn đề như thế nào?**

Như đã đề cập, các mạng nơ-ron sâu vượt trội trong việc đưa ra dự đoán dựa trên dữ liệu phần lớn không có cấu trúc. Điều đó có nghĩa là họ cung cấp hiệu suất tốt nhất trong các lĩnh vực như nhận dạng giọng nói và hình ảnh, nơi họ làm việc với dữ liệu lộn xộn như ghi âm lời nói và hình ảnh.

**Chúng ta có nên lúc nào cũng sử dụng Deep Learning thay vì Machine Learning?**

Không, bởi vì học sâu có thể rất tốn kém từ quan điểm tính toán.

Đối với các tác vụ không tầm thường, việc đào tạo một mạng lưới thần kinh sâu thường sẽ yêu cầu xử lý một lượng lớn dữ liệu bằng cách sử dụng các cụm GPU cao cấp trong nhiều, nhiều giờ.

Với các GPU hàng đầu có thể tốn hàng ngàn đô la để mua hoặc lên tới 5 đô la mỗi giờ để thuê trên đám mây, thật không khôn ngoan khi nhảy thẳng vào tìm hiểu sâu.

Nếu vấn đề có thể được giải quyết bằng thuật toán Machine Learning đơn giản hơn như suy luận Bayes hoặc hồi quy tuyến tính, thì không yêu cầu hệ thống phải vật lộn với sự kết hợp phức tạp của các tính năng phân cấp trong dữ liệu, thì các tùy chọn yêu cầu tính toán ít hơn này sẽ là sự lựa chọn tốt hơn

Deep Learning cũng có thể không phải là lựa chọn tốt nhất để đưa ra dự đoán dựa trên dữ liệu. Ví dụ: nếu tập dữ liệu nhỏ thì đôi khi các mô hình Machine Learning tuyến tính đơn giản có thể mang lại kết quả chính xác hơn – mặc dù một số chuyên gia về Machine Learning cho rằng mạng thần kinh Deep Learning được đào tạo đúng cách vẫn có thể hoạt động tốt với một lượng nhỏ dữ liệu.

AI và sức khỏe: Sử dụng Machine Learning để hiểu hệ thống miễnLearning dịch của con người  
HPE sẽ là kết thúc có hậu của MapR chứ?  
Microsoft: Nếu PC Windows 10 của bạn được AI chọn để cập nhật, sự cố sẽ ít xảy ra hơn  
Các nhà nghiên cứu của Nvidia sử dụng học tập sâu để tạo ra các video chuyển động siêu chậm

**Một trong những nhược điểm của Deep Learning là gì?**

Một trong những nhược điểm lớn là lượng dữ liệu họ cần đào tạo, gần đây Facebook tuyên bố họ đã sử dụng một tỷ hình ảnh để đạt được hiệu suất phá kỷ lục bởi một hệ thống nhận dạng hình ảnh. Khi các bộ dữ liệu lớn như vậy, các hệ thống đào tạo cũng yêu cầu quyền truy cập vào một lượng lớn sức mạnh tính toán phân tán. Đây là một vấn đề khác của học tập sâu, chi phí đào tạo. Do kích thước của bộ dữ liệu và số chu kỳ đào tạo phải được chạy, đào tạo thường yêu cầu quyền truy cập vào phần cứng máy tính mạnh mẽ và đắt tiền, điển hình là GPU cao cấp hoặc mảng GPU. Cho dù bạn đang xây dựng hệ thống của riêng mình hoặc thuê phần cứng từ nền tảng đám mây, không có tùy chọn nào có thể rẻ.

Mạng lưới thần kinh sâu cũng khó đào tạo, do cái được gọi là vấn đề độ dốc biến mất, có thể làm xấu đi nhiều lớp hơn trong mạng lưới thần kinh. Khi nhiều lớp được thêm vào, vấn đề độ dốc biến mất có thể dẫn đến việc mất một thời gian dài không thể để đào tạo một mạng lưới thần kinh đến một mức độ chính xác tốt, vì sự cải thiện giữa mỗi chu kỳ đào tạo là rất ít. Vấn đề không ảnh hưởng đến tất cả các mạng thần kinh nhiều lớp, thay vào đó là các mạng sử dụng phương pháp học tập dựa trên độ dốc. Điều đó nói rằng vấn đề này có thể được giải quyết theo nhiều cách khác nhau, bằng cách chọn một chức năng kích hoạt phù hợp hoặc bằng cách đào tạo một hệ thống sử dụng GPU hạng nặng.

Deep Learning giúp Google theo dõi nguy cơ đau tim  
Tương lai của tương lai: Spark, hiểu biết dữ liệu lớn, phát trực tuyến và Deep Learning trong đám mây

**Tại sao rất khó để đào tạo mạng lưới thần kinh Deep Learning?**

Như đã đề cập, mạng lưới thần kinh sâu rất khó đào tạo vì số lượng các lớp trong mạng lưới thần kinh. Số lượng các lớp và liên kết giữa các nơ-ron trong mạng sao cho khó có thể tính toán các điều chỉnh cần thực hiện ở mỗi bước trong quy trình đào tạo – một vấn đề được gọi là vấn đề độ dốc biến mất.

Một vấn đề lớn khác là số lượng lớn dữ liệu cần thiết để đào tạo mạng lưới thần kinh học tập sâu, với các tập huấn luyện thường đo kích thước petabyte.

**Có những kỹ thuật Deep Learning nào?**

Có nhiều loại mạng lưới thần kinh sâu, với các cấu trúc phù hợp với các loại nhiệm vụ khác nhau. Ví dụ: Mạng thần kinh chuyển đổi (CNN) thường được sử dụng cho các tác vụ thị giác máy tính, trong khi Mạng thần kinh tái phát (RNN) thường được sử dụng để xử lý ngôn ngữ. Mỗi lớp có các chuyên môn riêng, trong CNN, các lớp ban đầu được chuyên biệt để trích xuất các tính năng riêng biệt từ hình ảnh, sau đó được đưa vào mạng thần kinh thông thường hơn để cho phép hình ảnh được phân loại. Trong khi đó, RNN khác với mạng nơ ron chuyển tiếp thức ăn truyền thống ở chỗ chúng không chỉ cung cấp dữ liệu từ lớp thần kinh này sang lớp thần kinh tiếp theo mà còn có các vòng phản hồi tích hợp, trong đó đầu ra dữ liệu từ một lớp được đưa trở lại lớp trước nó – cho mạng một dạng bộ nhớ. Có một dạng RNN chuyên biệt hơn bao gồm cái được gọi là ô nhớ và được điều chỉnh để xử lý dữ liệu có độ trễ giữa các đầu vào.

Loại mạng thần kinh cơ bản nhất là mạng perceptron nhiều lớp, loại được thảo luận ở trên trong ví dụ về các số liệu viết tay, trong đó dữ liệu được đưa về phía trước giữa các lớp tế bào thần kinh. Mỗi nơ-ron thường sẽ biến đổi các giá trị mà chúng được cung cấp bằng cách sử dụng chức năng kích hoạt, thay đổi các giá trị đó thành một dạng, ở cuối chu kỳ đào tạo, sẽ cho phép mạng tính toán được bao xa để đưa ra dự đoán chính xác.

Có một số lượng lớn các loại mạng thần kinh sâu khác nhau. Không có một mạng nào tốt hơn mạng kia, chúng chỉ phù hợp hơn để học các loại nhiệm vụ cụ thể.

Gần đây, các mạng đối nghịch chung (Gans) đang mở rộng những gì có thể để sử dụng các mạng thần kinh. Trong kiến ​​trúc này, hai mạng thần kinh chiến đấu, mạng máy phát điện cố gắng tạo ra dữ liệu “giả” thuyết phục và người phân biệt đối xử cố gắng phân biệt sự khác biệt giữa dữ liệu giả và dữ liệu thực. Với mỗi chu kỳ đào tạo, máy phát điện trở nên tốt hơn trong việc tạo ra dữ liệu giả và người phân biệt đối xử có được con mắt sắc nét hơn để phát hiện ra những giả mạo đó. Bằng cách kết hợp hai mạng với nhau trong quá trình đào tạo, cả hai có thể đạt được hiệu suất tốt hơn. GAN đã được sử dụng để thực hiện một số nhiệm vụ quan trọng.