**TRƯỜNG ĐẠI HỌC BÁCH KHOA**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN LẬP TRÌNH TÍNH TOÁN**

**Đề tài 503: Bài toán dự đoán doanh số bán hàng sử dụng mạng neural**

Người hướng dẫn**: PGS. TS. NGUYỄN TẤN KHÔI**

Sinh viên thực hiện**:**

**Tên sinh viên 1: Nguyễn Trí Hoài Thương**

**Tên sinh viên 2: Hoàng Thị Hồng Thắm**

**LỚP: 21T-DT2 NHÓM: 21Nh10**

**Đà Nẵng, 04/2022**

MỤC LỤC

[MỤC LỤC i](#_Toc37650459)

[DANH MỤC HÌNH VẼ ii](#_Toc37650460)

[MỞ ĐẦU {font: TimeNew Roman, bold, size: 14, căn lề: center} i](#_Toc37650461)

[1. TỔNG QUAN ĐỀ TÀI 1](#_Toc37650462)

[2. CƠ SỞ LÝ THUYẾT 1](#_Toc37650463)

[2.1. Ý tưởng 1](#_Toc37650464)

[2.2. Cơ sở lý thuyết 1](#_Toc37650465)

[3. TỔ CHỨC CẤU TRÚC DỮ LIỆU VÀ THUẬT TOÁN 1](#_Toc37650466)

[3.1. Phát biểu bài toán 1](#_Toc37650467)

[3.2. Cấu trúc dữ liệu 1](#_Toc37650468)

[3.3. Thuật toán 1](#_Toc37650469)

[4. CHƯƠNG TRÌNH VÀ KẾT QUẢ 1](#_Toc37650470)

[4.1. Tổ chức chương trình 1](#_Toc37650471)

[4.2. Ngôn ngữ cài đặt 1](#_Toc37650472)

[4.3. Kết quả 1](#_Toc37650473)

[4.3.1. Giao diện chính của chương trình 1](#_Toc37650474)

[4.3.2. Kết quả thực thi của chương trình 1](#_Toc37650475)

[4.3.3. Nhận xét đánh giá 1](#_Toc37650476)

[5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 1](#_Toc37650477)

[5.1. Kết luận 1](#_Toc37650478)

[5.2. Hướng phát triển 1](#_Toc37650479)

[TÀI LIỆU THAM KHẢO 2](#_Toc37650480)

DANH MỤC HÌNH VẼ

No table of figures entries found.

MỞ ĐẦU{font: TimeNew Roman, bold, size: 14, căn lề: center}

{Để 2 dòng trống}

{Font: Time New Roman; thường; cỡ chữ: 13; dãn dòng: 1,3; căn lề: justified}

{Trong phần này, cần trình bày về: Mục đích thực hiện đề tài, mục tiêu đề tài, phạm vi và đối tượng nghiên cứu, phương pháp nghiên cứu, cấu trúc của đồ môn học}

# Lý do chọn đề tài:

Trong những năm qua, tầm quan trọng của máy tính trong việc lưu trữ và xử lý thông tin ngày càng được quan tâm. Các thiết bị thu thập dữ liệu tự động cũng không ngừng phát triển góp phần tạo ra cơ sở dữ liệu khổng lồ. Nhưng để có những thông tin bổ ích cho việc ra quyết định cũng như dự báo các chỉ số trong tương lai góp phần phát huy tối đa nguồn “tri thức” to lớn. Với những yêu cầu đó, các mô hình cơ sở dữ liệu truyền thống và ngôn ngữ thao tác không còn phù hợp nữa. Để có thể khải phá cơ sở dữ liệu biến nó thành “tri thức” người ta đã nghiên cứu về tổ chức các kho dữ liệu và kho thông tin, các hệ trợ giúp ra quyết định, các phương pháp khai phá dữ liệu và phát hiện tri thức trong cơ sở dữ liệu. Đề tài tập trung nghiên cứu kỹ thuật sử dụng mạng nơron để đưa ra dự đoán giúp người dùng trong việc đưa ra quyết định. Trên cơ sở đó, đề tài nghiên cứu xây dựng chương trình dự báo doanh số bán hàng sử dụng mạng nơron lan truyền ngược.

# Mục tiêu và nhiệm vụ chính:

Tìm hiểu các đặc trưng của mạng nơ-ron nhân tạo, khả năng và các nguyên tắc để ứng dụng thành công mạng nơ-ron nhân tạo trong thực tế. Nghiên cứu ứng dụng mạng nơ-ron nhân tạo vào lớp bài toán dự đoán và phân loại.

Xây dựng chương trình “dự đoán doanh số bán hàng” nhằm phân tích, dự báo, hỗ trợ cho những người quan tâm.

# Đối tượng và phạm vi nghiên cứu:

Đối tượng nghiên cứu là lớp bài toán dự đoán và phân loại, sử dụng mạng nơ-ron nhân tạo truyền thẳng huấn luyện bằng thuật toán lan truyền ngược.

Phạm vi nghiên cứu là lý thuyết ứng dụng mạng nơ-ron nhân tạo cho bài toán dự đoán và phân loại.

# Phương pháp nghiên cứu:

Phương pháp nghiên cứu tài liệu: nghiên cứu lý thuyết và ứng dụng mạng nơron nhân tạo trong bài toán dự đoán và phân loại.

Phương pháp thực nghiệm: nghiên cứu ứng dụng mạng nơ-ron nhân tạo bắt đầu từ bước chuẩn bị dữ liệu. Thực hiện phân tích ứng dụng mạng nơ-ron vào bài toán dự đoán. Từ các phân tích bài toán, xây dựng quy trình, chương trình để thực hiện, mô phỏng mạng nơ-ron dự báo doanh số bán hàng.

# Ý nghĩa khoa học và thực tiễn của đề tài:

Đề tài làm rõ khả năng ứng dụng của mạng nơ-ron trong bài toán dự đoán và phân loại. Xây dựng thành quy trình với các bước thực hiện cụ thể cho việc giải bài toán dự đoán và phân loại bằng mạng nơ-ron.

# Bố cục đồ án môn học:

Toàn bộ nội dung đồ án được trình bày trong 3 chương:

Chương 1: Cơ sở lý thuyết.

Chương 2: Tổ chức cấu trúc dữ liệu và thuật toán.

Chương 3: Chương trình và kết quả.

Chương 4: Kết luận và hướng phát triền.

# TỔNG QUAN ĐỀ TÀI

{Nội dung A {Font: Time New Roman; thường; cỡ chữ: 13; dãn dòng: 1,3; căn lề: justified}

Tìm hiểu và nắm vững cơ sở lý thuyết về kỹ thuật mạng Nơron nhân tạo, thuật giải di truyền, kinh tế học, phân tích tác động chính sách kinh tế.

Vận dụng xây dựng chương trình “Dự báo doanh số bán hàng”. Công cụ nhằm phân tích,dự báo, hỗ trợ cho những người quan tâm và cơ quan có chức năng.

# CƠ SỞ LÝ THUYẾT

## Ý tưởng

Tìm hiểu các đặc trưng của mạng nơ-ron nhân tạo, khả năng và các nguyên tắc để ứng dụng thành công mạng nơ-ron nhân tạo trong thực tế. Nghiên cứu ứng dụng mạng nơ-ron nhân tạo vào lớp bài toán dự đoán và phân loại.

Xây dựng chương trình “dự đoán doanh số bán hàng” nhằm phân tích, dự báo, hỗ trợ cho những người quan tâm.



*Hình 1. Neutral Networks*

## Cơ sở lý thuyết:

### Mô hình dự báo:

* **Mô hình ứng dụng mạng nơron**: Là mô hình có khả năng “học” từ các dữ liệu quá khứ, có thể cập nhật các tham số. Nếu lựa chọn được các tham số tối ưu thì đó là mô hình xấp xỉ rất tốt đường cong dịch chuyển của đối tượng cần dự báo. Kết quả cũng có độ lệch chính xác cao

### D:\NghienCuuKhoaHoc\noronsinhhoc.jpgMạng nơ-ron nhân tạo:

*Hình 2: Mô hình mạng nơron sinh học*

Mạng nơ-ron sinh học là một mạng lưới (plexus) các nơ-ron có kết nối hoặc có liên quan về mặt chức năng trực thuộc hệ thần kinh ngoại biên (peripheral nervous system) hay hệ thần kinh trung ương (central nervous system). Trong ngành thần kinh học (neuroscience), nó thường được dùng để chỉ một nhóm nơ-ron thuộc hệ thần kinh là đối tượng của một nghiên cứu khoa học nhất định.

Hệ thần kinh con người có khoảng 1010 tế bào thần kinh được gọi là các nơron, mỗi nơron có thể liên kết với 104 nơron khác thông qua khớp nối.

Mỗi nơron gồm có 3 phần: thân nơron có nhiệm vụ tiếp nhận hay phát ra các xung thần kinh, bên trong có nhân (Soma), hệ thống dây thần kinh vào (dendrites – còn gọi là các nhánh thụ giác) và một đầu dây thần kinh ra (sợi trục axon – nhánh trực giác) để dẫn truyền các xung thần kinh. Các đầu dây thần kinh vào nhận tín hiệu từ các nơron khác, nhân nơron sẽ sinh ra tín hiệu ở đầu ra của nơron và truyền tới các nơron khác được kết nối với đầu ra qua trục.

#### Quá trình phát triển, mô hình và quá trình xử lý trong mạng nơron nhân tạo.

##### Quá trình phát triển

Sự phát triển của mạng nơron trải qua cả quá trình đưa ra các khái niệm mới lẫn thực thi khái niệm này. Dưới đây là các mốc đáng chú ý trong lịch sử phát triên mạng nơron:

Cuối thế kỉ 19, đầu thế kỉ 20, sự phát triển chủ yếu chỉ là các công việc có sự tham gia của cả ba ngành Vật lý học, Tâm lý học, và Thần kinh học bởi các nhà khoa học như Hemholtz, Ernst Mach, Ivan Pavlov. Các công trình nghiên cứu của họ chủ yếu đi sâu vào các lý thuyết tổng quát về HỌC (learning), NHÌN (vision), và LẬP LUẬN (conditioning); và không hề đưa ra những mô hình toán học cụ thể mô tả hoạt động của các nơron.

Đầu những năm 1940, hai nhà khoa học Warren McCulloch và Walter Pitts đã chỉ ra rằng về nguyên tắc, mạng của các nơron nhân tạo có thể tính toán bất kì một hàm số học hay logic nào.

Đến cuối những năm 1950, ứng dụng thực nghiệm đầu tiên của mạng nơron nhân tạo được ra đời với phát minh của mạng nhận thức (perceptron network) và luật học tương ứng bởi Frank Rosenblatt. MẠng này có khả năng nhận thức các mẫu điều này mở ra hy vọng cho việc nghiên cứu mạng nơron nhân tạo.

Cùng thời gian đó, Bernard Widrow và Ted Hoff đã đưa ra một thuật toán mới và được sử dụng nó huấn luyện mạng nơron tuyến tính thích nghi. Luật học này vẫn được sử dụng cho đến ngày nay.

Tuy nhiên cả Bernard Widrow và Ted Hoff đều gặp phải trở ngại do Marvin Minsky và Seymour phát hiện ra. Đó là các mạng nhận thức chỉ có khả năng giải quyết được các bài toán khả phân tuyến tính. Họ đã cố gắng cải tiến nhưng không thành công do không có được các máy tinh đủ mạnh để có thể thực nghiệm.

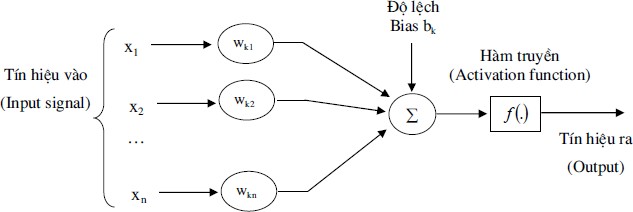
Cho đến nhăng năm 1980 việc nghiên cứu mạng nơron phát triển mạnh mẽ cùng với sự phát triển vượt bậc của PC. Có hai khái niệm mới liên quan tới sự hồi sinh này:

* + - * + Việc sử dụng các phương pháp thống kê để giải thích hoạt động của một lớp các mạng hồi quy (recurrent network) có thể được dùng như bộ nhớ liên hợp (associative memory) trong công trình nghiên cứu nhà vật lý học Johh Hopfield.
        + Sự ra đời của thuật toán lan truyền ngược (back - propagation) để luyện các mạng nhiều lớp.

##### Mô hình và quá trình xử lý trong mạng nơron nhân tạo

Để mô phỏng các tế bào thần kinh và các khớp nối thần kinh của bộ não con người, trong mạng nơ-ron nhân tạo củng có các thành phần có vai trò tương tự là các nơ-ron nhân tạo và kết nối giữa chúng( kết nối này gọi là weights).

Nơron là một đơn vị tính toán có nhiều đầu vào và một đầu ra, mỗi đầu vào đến từ một syanpse. Đặc trưng của nơron là một hàm kích hoạt phi tuyến chuyển đổi một tổ hợp tuyến tính của tất cả các tín hiệu đầu vào thành tín hiệu đầu ra.

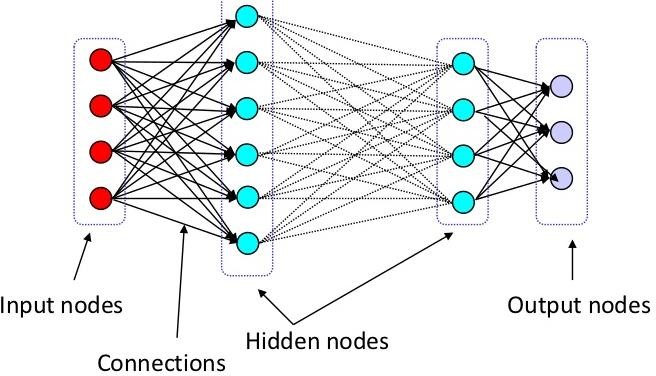


*Hình 2: Mô hình của một mạng nơron nhân tạo*

* Đầu vào của nơron gồm n tín hiệu x = (𝑥1, 𝑥2, …, 𝑥𝑛), đầu ra tín hiệu y = (𝑦1, 𝑦2, …, 𝑦𝑚).
* Một tập hợp các khớp nối và trọng số tương ứng với 𝑤𝑖.
* Một bộ cộng ∑ thực hiện trên các trọng số các khớp nối thường được gọi là bộ kết hợp tính tuyến.
* Một hàm chuẩn khống chế giá trị đầu ra của mạng nơron được gọi là hàm truyền hay hàm kích hoạt. Thông thường, tín hiệu đầu ra của một nơron trong khoảng [0,1] hoặc [-1, 1].
* Trang thái bên trong của nơron được xác định qua bộ tổng các đầu vào có trọng số w (i = 1, 2, …, n). Đầu ra y được xác định qua hàm phi tuyến f.

#### Cấu trúc và phân loại mạng nơron nhân tạo

##### Cấu trúc của mạng nơron nhân tạo

Trong mô hình mạng nơron nhân tạo, các nơron được nối với nhau bằng các liên kết nơron, mỗi lien kết có một trọng số đặc trưng cho đặc tính kích hoạt hay ức chế giữa các nơron. Đồng thời, các nơron được nhóm lại với nhau theo cấu trúc phân lớp, bao gồm: lớp vào (input layer), lớp ra (output layer) và lớp ẩn (hidden layer).

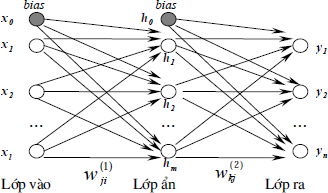
*Hình 3: Mô hình cấu trúc của mạng nơron nhân tạo*

* Lớp vào: Các nút trong lớp vào được gọi là các nút vào, chúng sẽ được mà hóa để đưa vào mạng xử lý. Các nơron vào không xử lý thông tin mà chỉ phán tán thông tin cho các nút khác.
* Lớp ẩn: Các nơron trong lớp ẩn chúng tạo thành các mô hình toán học phi tuyến cho mạng và chúng không thể nhìn thấy trực tiếp.
* Lớp ra: Các nơron này được gọi là các nút ra, nhiệm vụ của chúng là đưa thông tin ra thích nghi mẫu mã nguồn sử dụng cần.

##### Phân loại mạng nơron

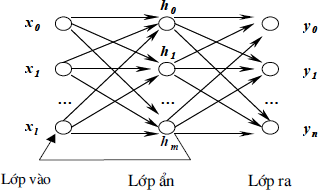
###### Phân loại theo liên kết:

* Mạng nơron truyền thẳng (feed – forward Neural Network): Trong mạng nơron này, các liên kết nơron chỉ truyền theo một hướng từ lớp vào cho tới lớp ra, không tạo thành các chu trình với các đỉnh là các nút nơron, các cung là các liên kết giữa chúng.



*Hình 4: Mạng nơron truyền thẳng*

* Mạng nơron hồi quy: Cho phép các liên kết nơron tạo thành chu trình, có thông tin xử lý hai chiều. Mạng hồi quy có khả năng lưu trữ trạng thái trong của nó dưới dạng các ngưỡng kích hoạt ngoài các trọng số liên kết nơron vì các thông tin ra của các nơron được truyền lại cho các nơron đã góp phần kích hoạt chúng.



*Hình 5: Mạng nơron hồi quy*

* Mạng nơron đối xứng và bất đối xứng: Nếu thỏa mãn yêu cầu có một đường nối từ nút i đến nút j thì cũng có một đường nối từ nút j đến nút I và trọng số tương ứng với hai đường nối này bằng nhau: 𝑤𝑗𝑖 = 𝑤𝑖𝑗 . Mạng không thỏa mảng điều kiện trên là mạng bất đối xứng.

###### Phân loại theo số lớp:

* Mạng chỉ gồm một lớp vào và một lớp ra được gọi là mạng đơn hay mạng một lớp.
* Mạng có từ một lớp ẩn trở lên được gọi là mạng đa lớp hay mạng nhiều lớp.

Trong mô hình mạng đa lớp, đầu ra của các phần tử tính toán tại một lớp là đầu vào của lớp tiếp theo. Không cho phép các liên kết giữa các nơron trong cùng một lớp và các liên kết nơron nhảy qua một lớp trở lên.

#### Học và lan truyền trong mạng nơron nhân tạo

Mạng nơron thực hiện hai chức chính là học và tổng quát hóa. Học là quá trình hiệu chỉnh các tham số và trọng số liên kết trong mạng để tối thiểu hóa sai số với vector đầu cho trước. Quá trình học dừng lại khi thỏa mãn một điều kiện nào đó ví dụ như khi các trọng số của mạng tạo ra lỗi đủ nhỏ giữa đầu ra mong đợi và kết quả đầu ra của mạng với đầu vào

cho trước. Tổng quát hóa là quá trình đưa vào một vector đầu vào mới và sản sinh ra quyết định dựa trên vector đầu ra tính được từ mạng.

* Có hai kiểu học đó là: Học có tham số và Học cấu trúc

***Học có tham số****:* là các tham số về trọng số cập nhật kết nối giữa các nơron.

***Học cấu trúc****:* trọng tâm là sự biến đổi cấu trúc của các mạng nơron gồm số lượng nút và các loại liên kết.

##### Học và tổng quát hóa

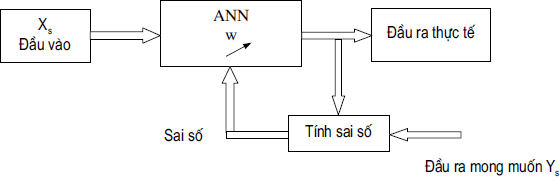
##### Học có giám sát

Học có giám sát là mạng được cung cấp một tập mẫu học {(𝑋𝑠, 𝑌𝑠 )} theo đó 𝑋𝑠

là tập tín hiệu đầu vào và kết quả đúng của hệ phải là 𝑌𝑠 . Ở mỗi lần học vector tín hiệu

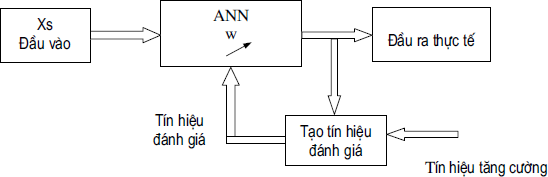
𝑋𝑠 được đưa vào mạng sao đó so sánh sự sai khác giữa các kết quả ra đúng 𝑌𝑠 với két ủa tính toán qua mạng. Quá trình tiếp tục cho đến khi thảo mãn một yêu cầu nào đó.

*Hình 6: Học có giám sát*



##### Học tăng cường

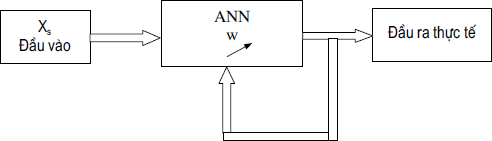
Cũng như kĩ thuật học giám sát, các vector đầu ra được biết chính xác. Nhưng trong trường hợp ít thông tin như output quá lớn hoặc chỉ đúng khoảng 40%. Khi đó một tín hiệu đánh giá là “True” hay “False” quay lại mạng.



*Hình 7: Học tăng cường*

##### Học không giám sát

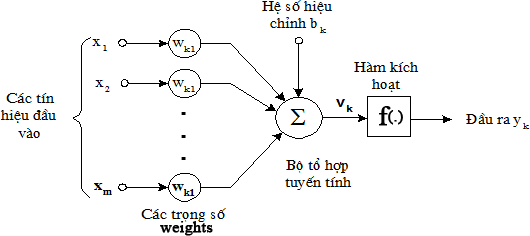
Trong mô hình học này, đầu ra mong muốn của mạng sẽ không được cho trước và mạng được trang bị khả năng tự tổ chức. Mạng không sử dụng mối quan hệ lớp của các mẫu học mà dùng thông tin kết hợp với nhóm các nơron để thay đổi các tham số cục bộ sao cho hợp nhất.



*Hình 8: Học không giám sát*

##### Lan truyền trong mạng nơron nhân tạo

###### Các định nghĩa và công thức



*Hình 9: Mô hình mạng nơron nhân tạo*

Ở đây, chúng ta xác định ba thành phần cơ bản của một mô hình nơron:

Một tập hợp các *weight*s hay *các kết n*ối, mà mỗi một trong chúng được đặc trưng bởi một trọng số của riêng nó. Tức là một tín hiệu 𝑥𝑗 tại đầu vào của weights j nối với nơron k sẽ được nhân với trọng số weights 𝑤𝑘𝑗 . Ở đó k là chỉ số của nơron tại đầu ra của weights đang xét, còn j chỉ điểm đầu vào của weights. Các trọng số weights cuả một nơron nhân tạo có thể nhận cả các giá trị âm và các giá trị dương.

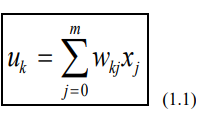
*Một bộ cộng* để tính tổng các tín hiệu đầu vào của nơron, đã được nhân với các trọng số weights tương ứng; phép toán được mô tả ở đây tạo nên một bộ tổ hợp tuyến tính.

*Một hàm kích hoạt (activation function)* để giới hạn biên độ đầu ra của nơron. Hàm kích hoạt cũng được xem xét như là một hàm nén; nó nén (giới hạn) phạm vi biên độ cho phép của tín hiệu đầu ra trong một khoảng giá trị hữu hạn. Mô hình nơron trong *hình 9* còn bao gồm một hệ số hiệu chỉnh

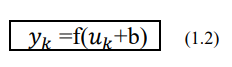
tác động từ bên ngoài, 𝑏𝑘. Hệ số hiệu chỉnh 𝑏𝑘 có tác dụng tăng lên hoặc giảm đi đầu vào thực của hàm kích hoạt, tuỳ theo nó dương hay âm.Có nhiều loại hàm kích hoạt.

Dưới dạng công thức toán học, chúng ta có thể mô tả một nơron *k* bằng cặp công thức

sau:



Và



ở đó 𝑥1, 𝑥2, ..., 𝑥𝑚 là các tín hiệu đầu vào; 𝑤𝑘1, 𝑤𝑘2,..., 𝑤𝑘𝑚là các trọng số weights của nơron *k*; 𝑢𝑘là đầu ra bộ tổ hợp tuyến tính tương ứng; 𝑏𝑘 là hệ số hiệu chỉnh.

Hệ số hiệu chỉnh 𝑏𝑘 là một tham số ngoài của nơron nhân tạo *k.* Chúng ta có thể thấy được sự có mặt của nó trong công thức (1.2). Một cách tương đương, chúng ta có thể tổ hợp các công thức (1.1) và (1.2) như sau:

*ek*   *wkj x j j* 0

*m*

(1.3)

Và

*yk*  *f* (*vk* )

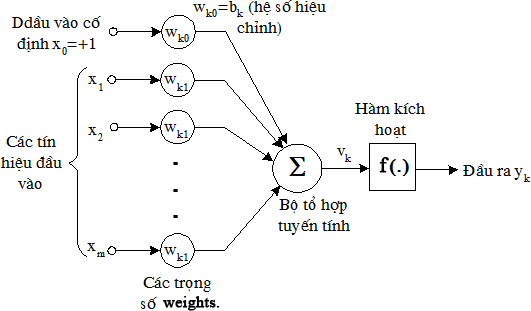
(1.4)

Trong công thức (1.3), chúng ta đã thêm một weights mới. Đầu vào của nó là:

𝐱𝟎**=+ 1**(1.5) và trọng số của nó là

𝐰𝐤𝟎**=** 𝐛𝐤(1.6)

Như vậy chúng ta vẽ lại mô hình của nơron *k* như trong hình (1.3). Trong hình này, nhiệm vụ của hệ số hiệu chỉnh là thực hiện hai việc: (1) thêm một tín hiệu đầu vào cố định là 1, và (2) thêm một trọng số weights mới bằng giá trị của hệ số 𝑏𝑘. Mặc dầu các mô hình trong hình (1.2) và (1.3) là khác nhau về hình thức nhưng tương tự về bản chất toán học.



*Hình 10: Mô hình phi tuyến thứ hai của một nơron*

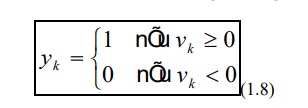
###### Các kiểu hàm kích hoạt

Hàm kích hoạt, ký hiệu bởi f(v), xác định đầu ra của nơron. Dưới đây là các kiểu hàm kích hoạt cơ bản:

ham nguong.PNG**1.Hàm ngưỡng:** Đối với loại hàm này, chúng ta có:

(1.7)

Trong các tài liệu kỹ thuật, dạng hàm ngưỡng này thường được gọi là hàm Heaviside. Đầu ra của nơron k sử dụng hàm ngưỡng sẽ như sau:



ở đó 𝑣𝑘 là đầu ra của bộ tổ hợp tuyến tính, có nghĩa là:

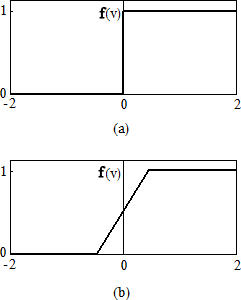
*vk*   *wkj x j*  *bk*

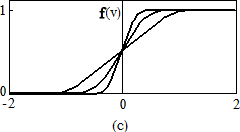
*j* 1

*m*

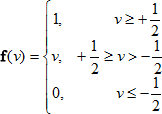
(1.9)

Một nơron như vậy thường được gọi là mô hình *McCulloch-Pitts*.





*Hình 11: (a): Hàm ngưỡng, (b): Hàm vùng tuyến tính, (c): Hàm sigma với tham số độ dốc a thay đổi*

**2.Hàm vùng tuyến tính:** Đối với loại hàm này (Hình 11.b)

(1.10)

Dạng hàm này có thể được xem như một xấp xỉ của một bộ khuếch đại phi tuyến

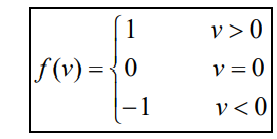
**3.Hàm sigma**

Hàm sigma là dạng chung nhất của hàm kích hoạt được sử dụng trong cấu trúc mạng nơron nhân tạo. Nó là một hàm tăng và nó thể hiện một sự trung gian giữa tuyến tính và phi tuyến. Một ví dụ của hàm này là *hàm logistics*, xác định như sau:

|  |  |
| --- | --- |
| f(v) = | 1 |
| 1+exp (−𝑎𝑣) |

Ở đó a là tham số độ dốc của hàm sigma. Bằng việc biến đổi tham số a, chúng ta thu được các hàm sigma với các độ dốc khác nhau, như được minh hoạ trong hình 11.c. Thực tế, hệ số góc tại v=0 là a/4. Khi tham số hệ số góc tiến tới không xác định, hàm sigma trở thành một hàm ngưỡng đơn giản. Trong khi một hàm ngưỡng chỉ có giá trị là 0 hoặc 1, thì một hàm sigma nhận các giá trị từ 0 tới 1. Cũng phải ghi nhận rằng hàm sigma là hàm phân biệt, trong khi hàm ngưỡng thì không (Tính phân biệt của hàm là một đặc tính quan trọng trong lý thuyết mạng nơron).

Các hàm kích hoạt được định nghĩa trong các công thức (1.7), (1.10), (1.11) đều trong phạm vi từ 0 đến 1. Đôi khi có những yêu cầu xây dựng hàm kích hoạt trong phạm vi từ -1 đến 1, trong trường hợp này hàm kích hoạt được giả định có dạng đối xứng qua gốc toạ độ (hay có thể gọi là dạng phản đối xứng); nghĩa là hàm kích hoạt là một hàm lẻ. Ví dụ, hàm ngưỡng trong công thức (1.7) bây giờ được xác định như sau:



( 1.12)

Hàm này thường được gọi là *hàm signum*.

Mạng nơron lan truyền thông tin từ lớp vào đến lớp ra. Khi việc lan truyền kết thúc, thông tin tại lớp ra chính là kết quả lan truyền.

# TỔ CHỨC CẤU TRÚC DỮ LIỆU VÀ THUẬT TOÁN

## Phát biểu bài toán

Mô tả đầu vào (Input) và đầu ra (Output)

**Đầu vào:** Một lớp tín hiệu vào

**Đầu ra:** Kết quả tương ứng với tập tín hiệu vào

* **Giải thuật:**

1. Đưa tập tín hiệu vào vào lớp vào của mạng.
2. Tính mức tích cực của các nút trong mạng.
3. Với mạng truyền thẳng: Nếu mức tích cực của nút ra đã biết thì kết thúc.

Với mạng phản hồi: Nếu mức tích cực của nút bằng hoặc xấp xỉ bằng hằng số thì kết thúc.

Nếu không thì quay lại bước 2.

**Giải thuật Back Propagaton:**

### Một số khái niệm cần biết:

Học trong mạng nơ-ron là quá trình thay đổi giá trị của w để mạng có thể đưa ra được kết quả mong muốn tối ưu. Sự khác nhau giữa giá trị thật của mẫu và kết quả dự đoán của mạng gọi là lỗi. Quá trình học sẽ tìm các wi\* để lỗi nhỏ nhất.

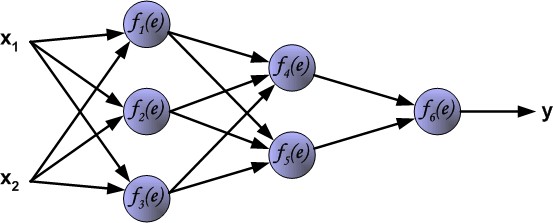
### Giải thuật:

Giải thuật Backpropagation gồm 2 giai đoạn:

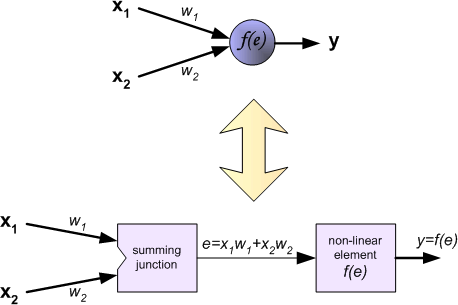
Lan truyền tiến (tính output của các nơ-ron).

Lan truyền ngược.

Để minh họa cho giải thuật , chúng ta sẻ xem xét giải thuật Backpropagation hiệu chỉnh các weights của một mạng nơ-ron gồm ba lớp với 2 đầu vào và một đầu ra .Như hình dưới:



Mỗi nơ-ron sẽ gồm hai thành phần chính. Thứ nhất là các trọng số weights và các trị số đầu vào x. Phần thứ 2 là hàm kích hoạt f. Và y = f(e) là đầu ra của một nơron.



Để dạy cho mạng chúng ta cần phải có các mẫu. Mỗi mẫu bao gồm hai thành phần đó là:

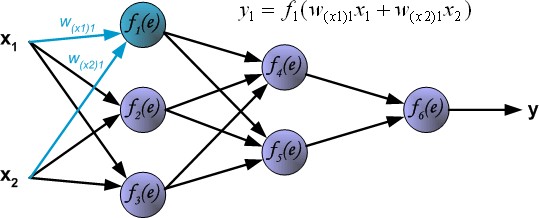
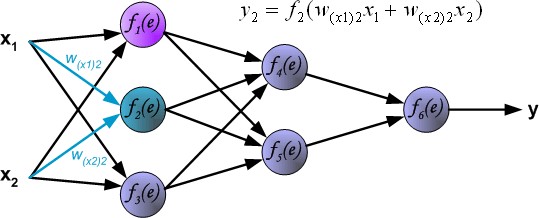
Đầu vào 𝑥1, 𝑥2.

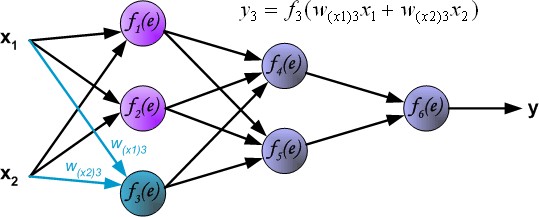
Đầu ra z là kết quả của mẫu.

#### Đầu tiên là bước lan truyền tiến:

Chúng ta sẽ bắt đầu với hai đầu vào 𝑥1, 𝑥2. Từ đó chúng ta có thể tính được đầu ra của mỗi nơron trong mạng. Hình dưới minh họa cho sự lan truyền các giá trị trong mạng.

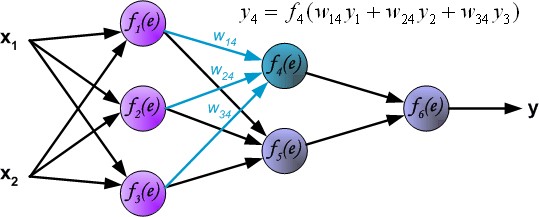
**Trong đó:**

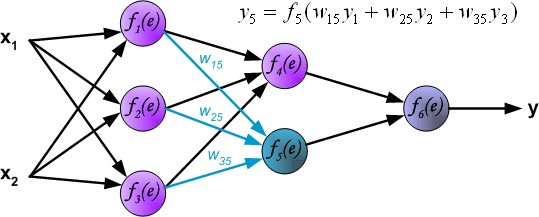
* W(xm)n: tượng trưng cho các trọng số weights của kết nối giữa đầu vào xm và nơron n trong tầng input của mạng.
* Yn tượng trưng cho đầu ra của nơron n.



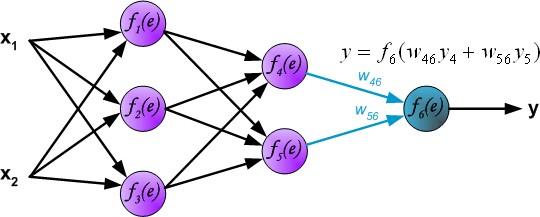
Sự lan truyền của các tín hiệu từ lớp đầu vào đến lớp ẩn.

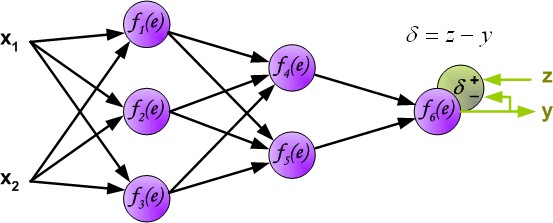
Kí hiệu wmn biểu diễn trọng số weights của kết nối giữa đầu ra của nơron m và là đầu vào của nơron n.





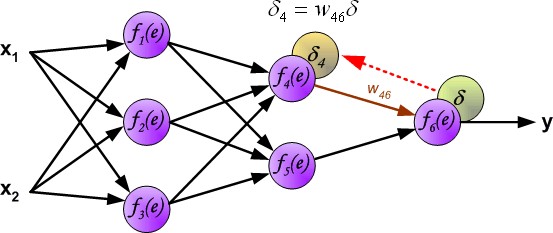
Sự lan truyền của tín hiệu đến lớp đầu ra:

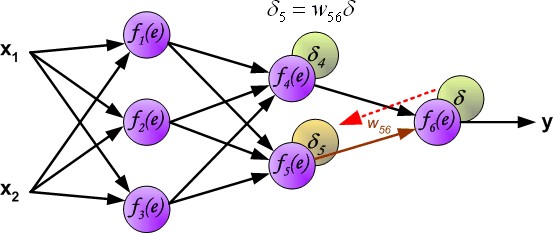


Bước tiếp theo của giải thuật là so sánh giá trị đầu ra của mạng là Y với giá trị z là kết quả của mẫu. Hiệu giữa chúng ta gọi là tín hiệu lỗi ** của lớp đầu ra.

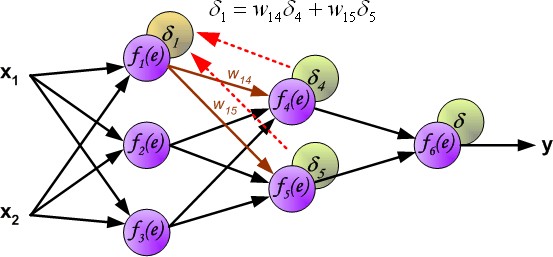
Chúng ta không thể tính được các tín hiệu lỗi của các nơron một cách trực tiếp. Bởi vì chúng ta không có được đầu ra của mỗi nơron. Nên chúng ta sẽ lan truyền tín hiệu lỗi ** lùi lại đến tất cả các nơron. ** n là tín hiệu lổi ở nơron n.

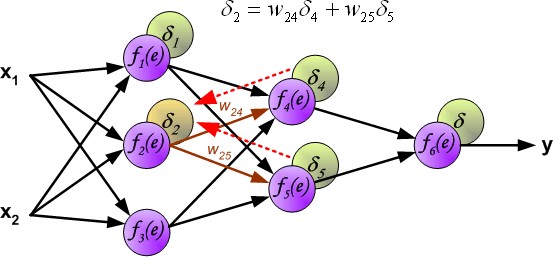
#### Lan truyền ngược:

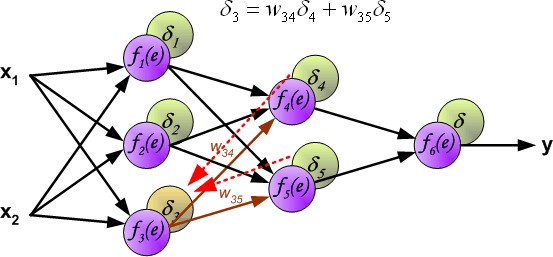


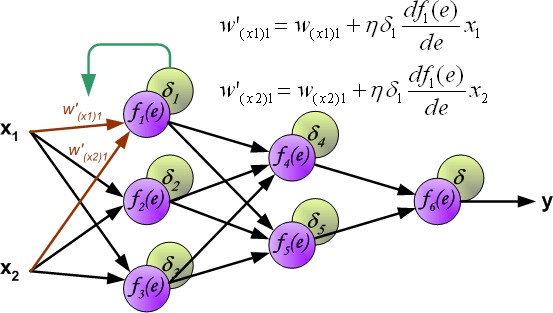


**Chú ý:** các trọng số weights wmn được dùng để tính toán trong trường hợp lan truyền ngược này bằng với các trọng số weights wmn trong lan truyền tiến . Chỉ có chiều của lan truyền tín hiệu là thay đổi. Kĩ thuật này được dùng cho tất cả các lớp của mạng. Ta có thể thấy bằng hình ảnh:



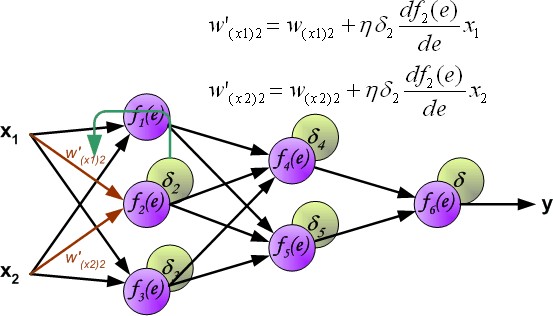


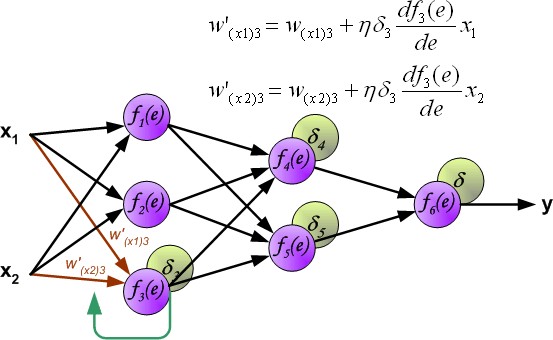


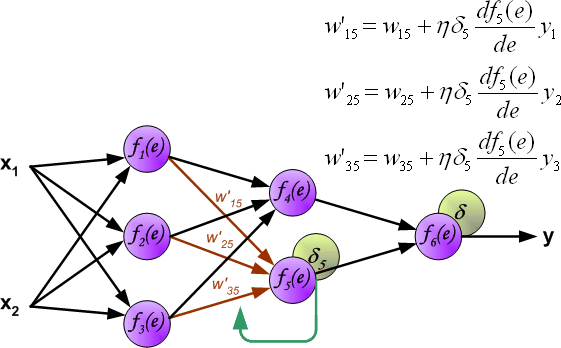
Khi tín hiệu lỗi của mỗi nơron đã được tính thì chúng ta có thể dựa vào tín hiệu lỗi này để thay đổi giá trị của các trọng số. Trong giải thuật Back Propagation trọng số thay đổi dựa vào đạo hàm của hàm kích hoạt trong mỗi nơron.

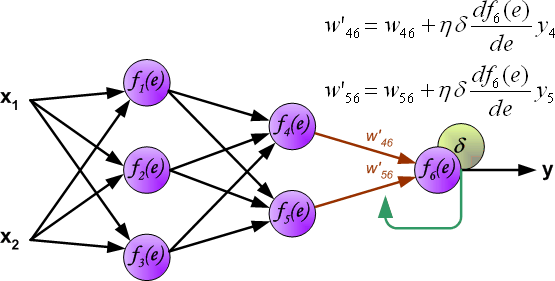
**Trong đó:**

* *df(e)/de* : là đạo hàm của hàm kích hoạt.
* ** là đại lượng ngẫu nhiên ảnh hưởng đến tốc độ dạy của mạng. Có nhiều cách để chọn

** Có thể bắt đầu quá trình học bằng cách chọn ** là một đại lượng lớn rồi trong suốt quá trình học sẽ giảm dần giá trị này. Hoặc cách khác nữa là bắt đầu bằng một giá trị nhỏ rồi sau đó trong quá trình học có thể tăng dần giá trị này, đến cuối quá trình học lại giảm giá trị này xuống.



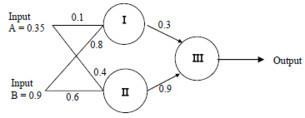




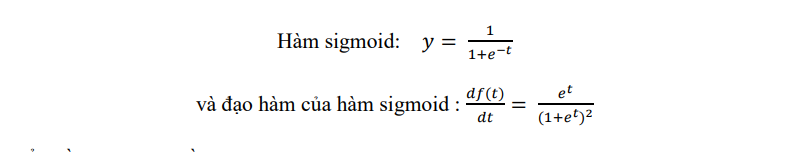
#### Ví dụ điển hình:

Sau đây chúng ta sẽ xét một ví dụ đơn giản về giải thuật này:

Giả sử chúng ta có một mạng nơron như sau:



Với mạng nơron cho trước các trọng số như hình với hàm kích hoạt là hàm sigmoid (sigma như đã nói ở trên), với hệ số hiệu chỉnh bias bằng 1 và hệ số * = 1*:



Mẫu gồm 3 thành phần:

Đầu vào: A = 0.35, B = 0.9.

Đầu ra mong muốn: z = 0.5.

Trước hết chúng ta sẽ tính đầu ra cho các nơron:

Nơ - ron 1 : Ta có : t1 = (0.35x0.1) + (0.9x0.8) + 1 = 1.755

Từ đó suy ra: y1 = = 0.855.

Nơ –ron 2 : ta có t2 = (0.9x0.6) + (0.35x0.4) + 1 = 1.68

Từ đó suy ra : y2 = = 0.842.

Nơ-ron 3: t3 = (0,3 x 0.855) + (0.9 x 0.842) + 1 = 2.0143

Y3 = = 0.8823

Bước tiếp theo là chúng ta sẻ tính tín hiệu lỗi

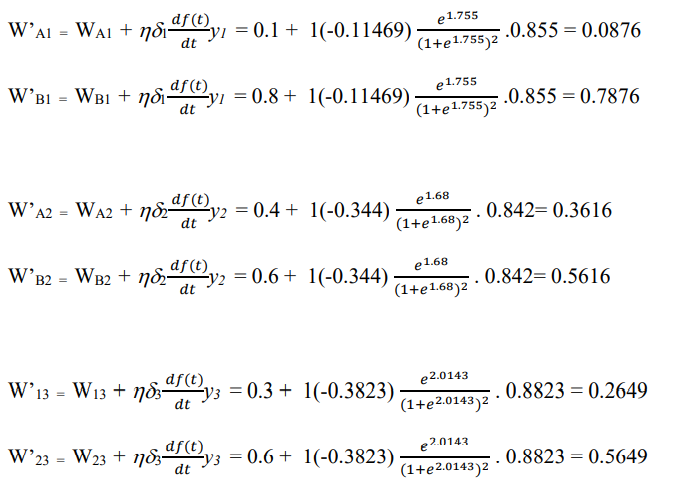
δ3 = z – y = 0.5 – 0.8823 = -0.3823.

Lan truyền tín hiệu lỗi:

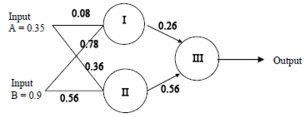
Nơ-ron 1: δ1 = −0.3823 \* 0.3 = −011469

Nơ-ron 2: δ2 = -0.3823 \* 0.9 = -0.344

Bây giờ chúng ta sẽ dựa vào tín hiệu lỗi của mổi nơron để thay đổi trọng số weights.



**Và đây là kết quả:**



## Cấu trúc dữ liệu

## Thuật toán

Trình bày các thuật toán và phân tích độ phức tạp của các thuật toán.

# CHƯƠNG TRÌNH VÀ KẾT QUẢ

## Tổ chức chương trình

## Cài đặt chương trình

## Kết quả thực hiện

### Giao diện chính của chương trình

### Các kết quả thực thi của chương trình

Mô tả kết quả 01 thực hiện chương trình, hình vẽ, giải thích.

Mô tả kết quả 02 thực hiện chương trình, hình vẽ, giải thích.

….

### Nhận xét đánh giá kết quả

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## Kết luận

## Hướng phát triển

TÀI LIỆU THAM KHẢO

PHỤ LỤC

Sinh viên bỏ Code từng phần vào đây.{*Font: Time New Roman; thường; cỡ chữ: 12; dãn dòng: 1,3; căn lề: justified}*