

## KHAI THÁC KẾT QUẢ HỌC TẬP SINH VIÊN BẰNG MẠNG NƠON NHÂN TẠO

Bùi Công Danh, Lư Nhật Vinh

*Trường Đại học Công nghiệp Thực phẩm TP.HCM*

Ngày gửi bài: 20/5/2016

Ngày chấp nhận đăng: 28/9/2016

### TÓM TẮT

Trong học chế tín chỉ quá trình chọn môn học tự chọn của sinh viên được xem như là một hệ thống với nhiều đầu vào và một đầu ra, trong đó các kết quả điểm tổng kết từng môn học trong quá khứ là đầu vào và kết quả điểm tổng kết của môn tự chọn là đầu ra. Bài báo trình bày về mạng nơon nhân tạo nhiều lớp với giải thuật học lan truyền ngược được áp dụng vào việc hỗ trợ sinh viên trong việc chọn môn học tự chọn sao phù hợp với năng lực của mình.

### DATA MINING LEARNING OUTCOMES OF STUDENT BY ARTIFICIAL NEURAL NETWORK

#### ABSTRACT

In the process of training subjects chosen by the students is considered as a system with multiple inputs and one output, which outlines the results for each subject in the past is the input and summarize points electives as output. This article presents an artificial neural network with multi-layer back-propagation learning algorithm is applied on assisting students in choosing electives stars suit their abilities.

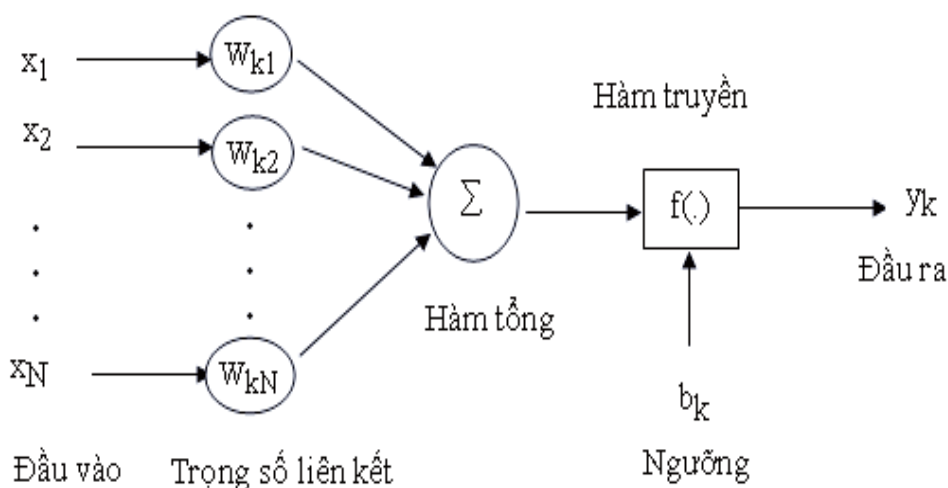
#### 1. GIỚI THIỆU

Trường Đại học Công nghiệp Thực phẩm Thành phố Hồ Chí Minh đã áp dụng hình thức học theo tín chỉ từ năm học 2006 đến nay. Theo tinh thần của học chế tín chỉ sinh viên có thể tự xây dựng kế hoạch học tập cho cả quá trình học của mình suốt khóa học. Có nghĩa là tùy theo năng lực và điều kiện của bản thân sinh viên sẽ đăng ký môn học nhiều hay ít trong mỗi học kỳ nhưng vẫn đảm bảo đúng thời hạn đào tạo theo qui định của Nhà trường trong bài báo này tác giả đề xuất mô hình tư vấn hỗ trợ sinh viên trong quá trình đăng ký môn học tự chọn sao cho phù hợp với điều kiện thời gian và năng lực nhằm tối ưu về mặt điểm tổng kết cho môn học tự chọn là một trong các boăn khoản của sinh viên trong việc đăng ký môn tự chọn.

Trong bài viết này, chúng tôi sẽ giới thiệu một cách tiếp cận sử dụng mạng nơon nhân tạo đã được áp dụng rộng rãi từ những năm 1980 trong các lĩnh vực như điều khiển thông minh, nhận dạng và xử lý tín hiệu. Một mạng nơon nhân tạo nhiều lớp với giải thuật học lan truyền ngược được áp dụng vào việc tư vấn chọn môn học tự chọn, mạng nơon này cho phép thiết lập một cầu nối giữa các thành phần điểm tổng kết của các môn học trong quá khứ và kết quả của môn học tự chọn trước khi sinh viên đăng ký học loại môn tự chọn này. Kết quả thực nghiệm bước đầu cho thấy mô hình tính khả quan của việc ứng dụng các phương pháp máy học vào hỗ trợ tư vấn chọn môn học tự chọn tại trường.

#### 2. MÔ HÌNH MẠNG NƠON NHÂN TẠO

Một nơon là một đơn vị xử lý thông tin và là thành phần cơ bản của một mạng nơon. Cấu trúc của một nơon được mô tả như hình dưới:



**Hình 1. Mô hình mạng nơron nhân tạo**

Các thành phần cơ bản của một nơron nhân tạo trong hình trên bao gồm:

- **Tập các đầu vào:** Là các tín hiệu vào (input signals) của nơron, các tín hiệu này thường được đưa vào dưới dạng một vector N chiều.
- **Tập các liên kết:** Mỗi liên kết được thể hiện bởi một trọng số (gọi là trọng số liên kết – Synaptic weight). Trọng số liên kết giữa tín hiệu vào thứ j với nơron k thường được kí hiệu là  $w_{kj}$ . Thông thường, các trọng số này được khởi tạo một cách ngẫu nhiên ở thời điểm khởi tạo mạng và được cập nhật liên tục trong quá trình học mạng.
- **Bộ tổng** (Summing function): Thường dùng để tính tổng của tích các đầu vào với trọng số liên kết của nó.
- **Ngưỡng** (còn gọi là một độ lệch - bias): Ngưỡng này thường được đưa vào như một thành phần của hàm truyền.
- **Hàm truyền** (Transfer function): Hàm này được dùng để giới hạn phạm vi đầu ra của mỗi nơron. Nó nhận đầu vào là kết quả của hàm tổng và ngưỡng đã cho. Thông thường, phạm vi đầu ra của mỗi nơron được giới hạn trong đoạn  $[0,1]$  hoặc  $[-1, 1]$ . Các hàm truyền rất đa dạng, có thể là các hàm tuyến tính hoặc phi tuyến. Việc lựa chọn hàm truyền nào là tùy thuộc vào từng bài toán và kinh nghiệm của người thiết kế mạng.
- **Đầu ra:** Là tín hiệu đầu ra của một nơron, với mỗi nơron sẽ có tối đa là một đầu ra.

### 3. BÀI TOÁN TƯ VẤN CHỌN MÔN HỌC TỰ CHỌN

Để chuẩn bị cho thời khóa biểu học của mình các sinh viên phải đăng ký môn học theo chương trình đào tạo đối với các môn bắt buộc thì sinh viên chỉ được lựa chọn buổi học, nhóm học, thứ học, hoặc tiết học.

Học kì 1			16	
Học phần bắt buộc			16	
1	19200001	Những nguyên lý cơ bản của chủ nghĩa Mác – Lênin	5(5,0,10)	
2	18200001	Toán cao cấp A1	2(2,0,4)	
3	18200013	Vật lý đại cương 1	2(2,0,4)	
4	21200001	Anh văn 1	2(2,0,4)	
5	01200001	Tin học đại cương	2(2,0,4)	
6	01201002	Thực hành tin học đại cương	1(0,1,1)	01200001(c)
7	19200004	Pháp luật đại cương	2(2,0,4)	
8	17201001	Giáo dục thể chất 1	60 tiết	
9	17200004	Giáo dục quốc phòng - an ninh 1	45 tiết	
10	17200005	Giáo dục quốc phòng - an ninh 2	45 tiết	

**Hình 2. Chương trình đào tạo học kỳ I ngành Công nghệ Thông tin**

Sau khi kết thúc mỗi môn học nhà trường thực hiện nhập điểm môn học vào phần mềm theo chương trình đào tạo của từng học kỳ.

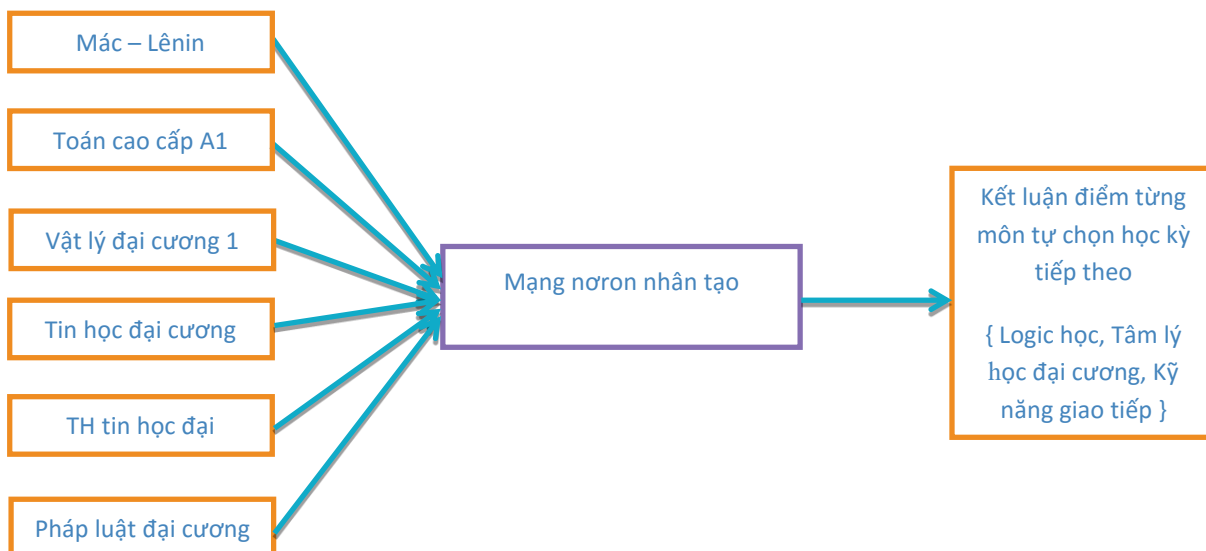
Ngày nay các chương trình đào tạo được thực hiện theo học chế tín chỉ nhằm tạo điều kiện thuận lợi cho quá trình học tập của sinh viên cũng như tăng cường khả năng tự học. Trong đó việc lựa chọn môn học phù hợp với thời gian, điều kiện đối với các môn bắt buộc theo chương trình đào tạo đã tạo ra một kết quả khả quan trong thời gian quan giúp sinh viên chủ động hơn trong việc học tập. Tuy nhiên bên cạnh đó cũng phát sinh một số băn khoăn trong việc đăng ký môn học tự chọn như: Chọn môn học nào để phù hợp với năng lực? Chọn môn học nào để được điểm cao? .v.v.

Học phần tự chọn			2	
1	18200016	Logic học	2(2,0,4)	
2	13200040	Tâm lý học đại cương	2(2,0,4)	
3	13200041	Kỹ năng giao tiếp	2(2,0,4)	

**Hình 3. Môn tự chọn trong một học kỳ**

#### 4. ỨNG DỤNG MẠNG NƠON NHÂN TẠO VÀO TƯ VẤN CHỌN MÔN TỰ CHỌN

Để hỗ trợ cho sinh viên trong việc chọn môn tự chọn khi đăng ký học phần, giảm bớt các boăn khoăn trong quá trình đăng ký môn học tự chọn. Tác giả đề xuất mô hình khai thác dữ liệu học tập của các học kỳ trước nhằm phân tích năng lực học tập của từng đối tượng sinh viên và đưa ra lời gợi ý thích hợp cho từng môn học tự chọn trước khi đăng ký loại môn này.



**Hình 4. Mô hình tư vấn chọn môn học**

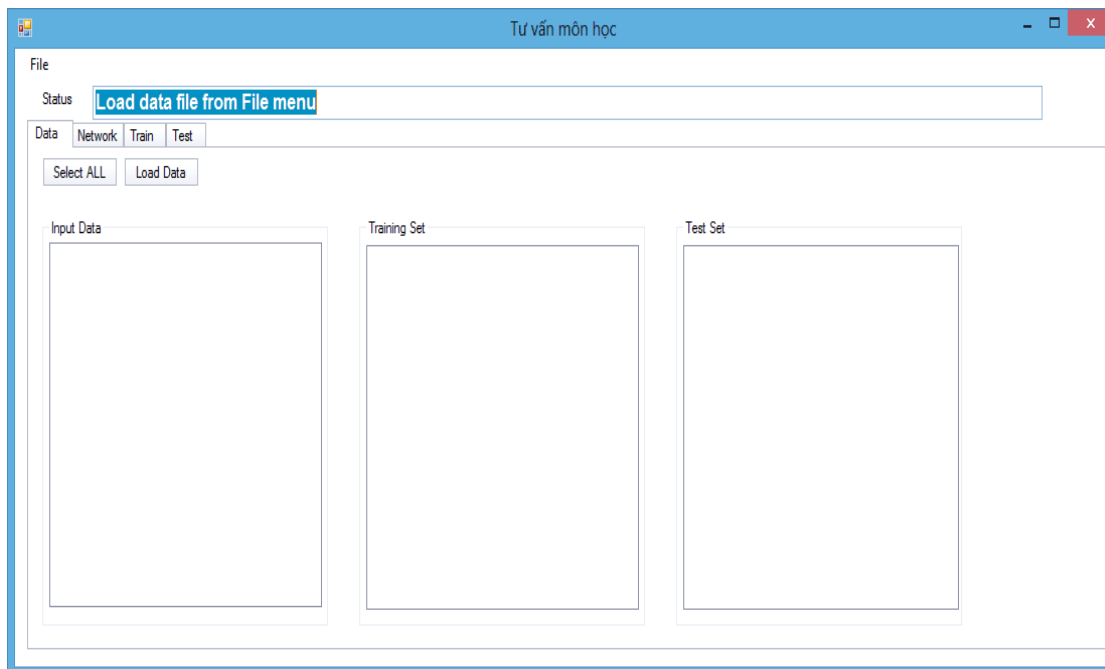
Cấu trúc tập tin sử dụng cho khai phá bằng mạng nơron.

**HuanLuyen.csv**: mỗi dòng chứa 7 thành phần dữ liệu của các môn học được sắp xếp như các vectơ và dữ liệu đầu ra phải được đặt trong cột cuối cùng tương tự như hình sau:

A	B	C	D	E	F	G	H	I	J
<ticker>	<MSSV>	<exchange>	<THDC>	<TH_THDC>	<T_A1>	<VL_A1>	<ML>	<PL>	<KNGT>
TCK.B	2001110002	T	6.3	8.5	7.2	5.1	6.1	6	5.7
TCK.B	2001110003	T	5.2	8.5	6.7	5.9	6	6.8	7.3
TCK.B	2001110008	T	5.1	8.5	6.1	4.6	5.7	6.2	6.8
TCK.B	2001110012	T	5.1	8	6.4	5.2	5.9	5.1	5.7
TCK.B	2001110013	T	5.7	8.5	6.3	5.9	6.1	6.9	5.5
TCK.B	2001110014	T	6.2	9	6.4	5.9	6.4	7.4	5.9
TCK.B	2001110016	T	6.1	8	5.8	4.2	5.4	6.1	5.1
TCK.B	2001110018	T	6.6	9	7.1	5.9	5.8	5.9	5.9
TCK.B	2001110020	T	5.5	7	5.6	5	6.4	5.6	5
TCK.B	2001110025	T	5.3	9	7.1	3.9	5.7	6.5	5.3
TCK.B	2001110026	T	3.8	9.3	6.2	3.6	6.4	6.6	5.9
TCK.B	2001110028	T	7.4	9.5	6	6	6.5	7.2	1
TCK.B	2001110033	T	8.4	10	7.4	5.8	6.4	7.4	5.5
TCK.B	2001110036	T	7.5	9.5	7.9	5.5	6.2	6.1	5.9
TCK.B	2001110037	T	3.8	8.5	5.7	4.7	5.4	5.6	5.5
TCK.B	2001110040	T	6.3	8	5.9	5.4	6.5	6.3	6.6
TCK.B	2001110042	T	5.8	8	5.2	5.2	6.6	6.6	6.6
TCK.B	2001110043	T	6.6	9.5	7.5	6.2	6.7	6.9	5.1
TCK.B	2001110045	T	6.6	9	7.2	5.4	6.4	6.9	3.2
TCK.B	2001110046	T	6.9	9	7.9	5.8	7.4	7.6	6.1
TCK.B	2001110047	T	6.4	9.8	2.7	5.6	6	6.7	3
TCK.B	2001110047	T	6.4	9.8	2.7	5.6	6	6.7	4.2
TCK.B	2001110047	T	6.4	9.8	2.7	5.6	6	6.7	3.2
TCK.B	2001110049	T	4.9	8	7.6	4.8	6.7	6.3	5.6

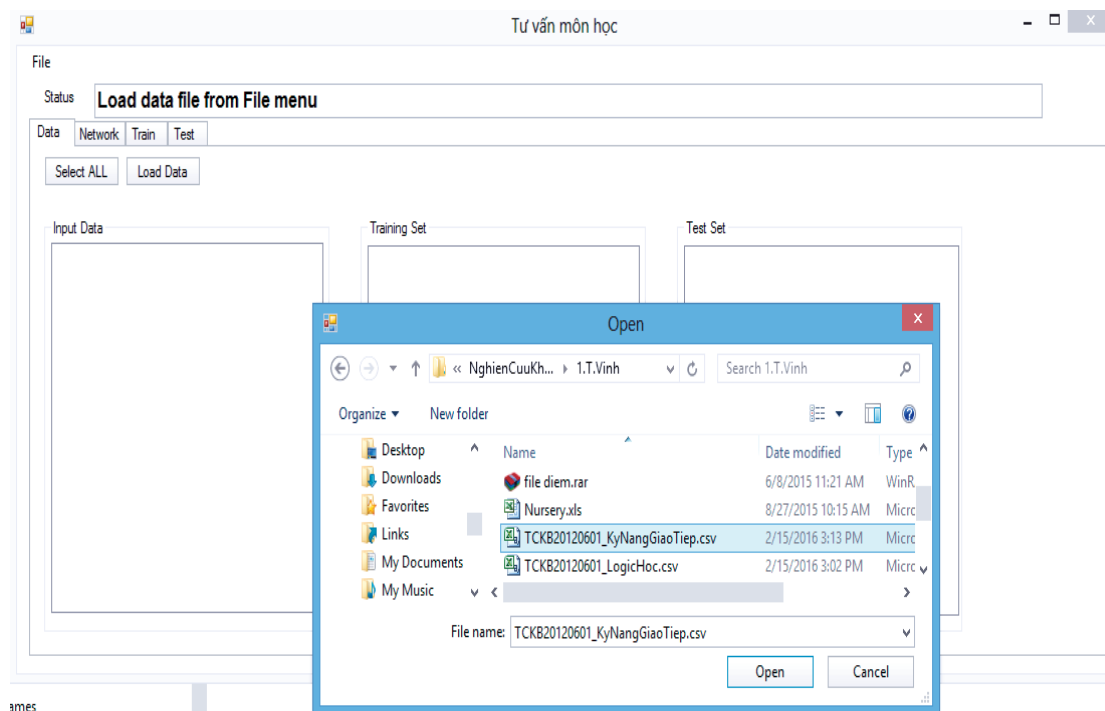
**Hình 5. Cấu trúc tập tin dữ liệu huấn luyện dùng mạng nơron**

### Kết quả hiện thực chương trình

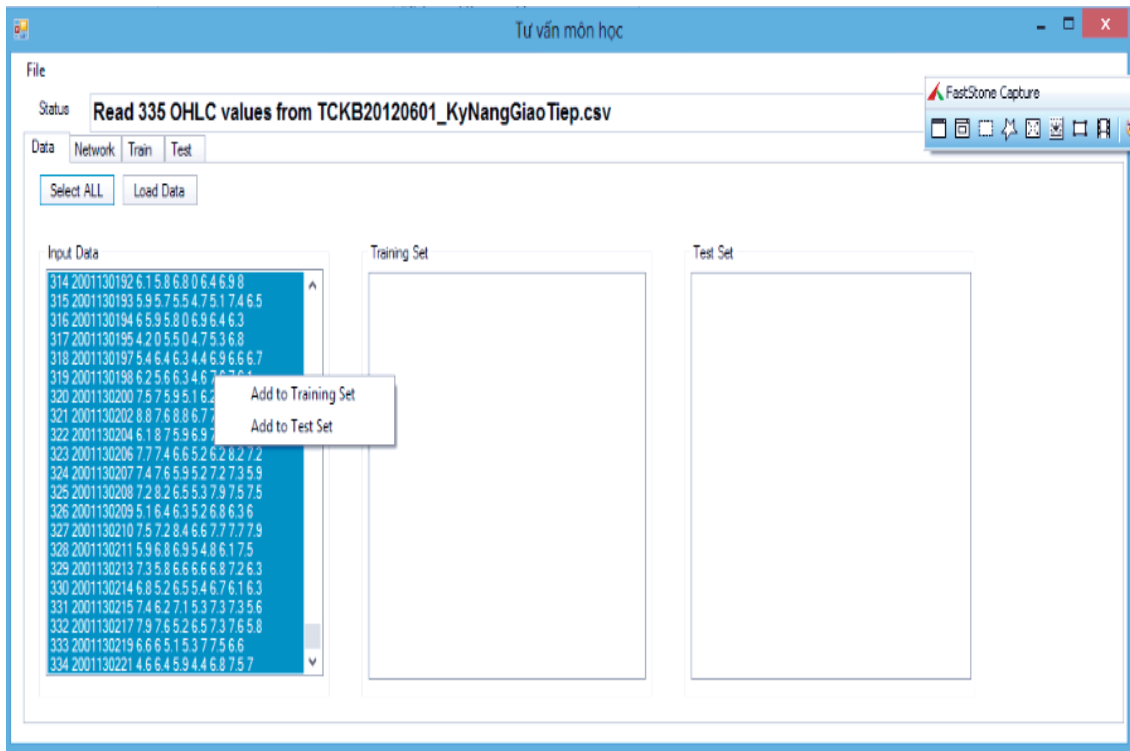


Hình 6. Màn hình chính

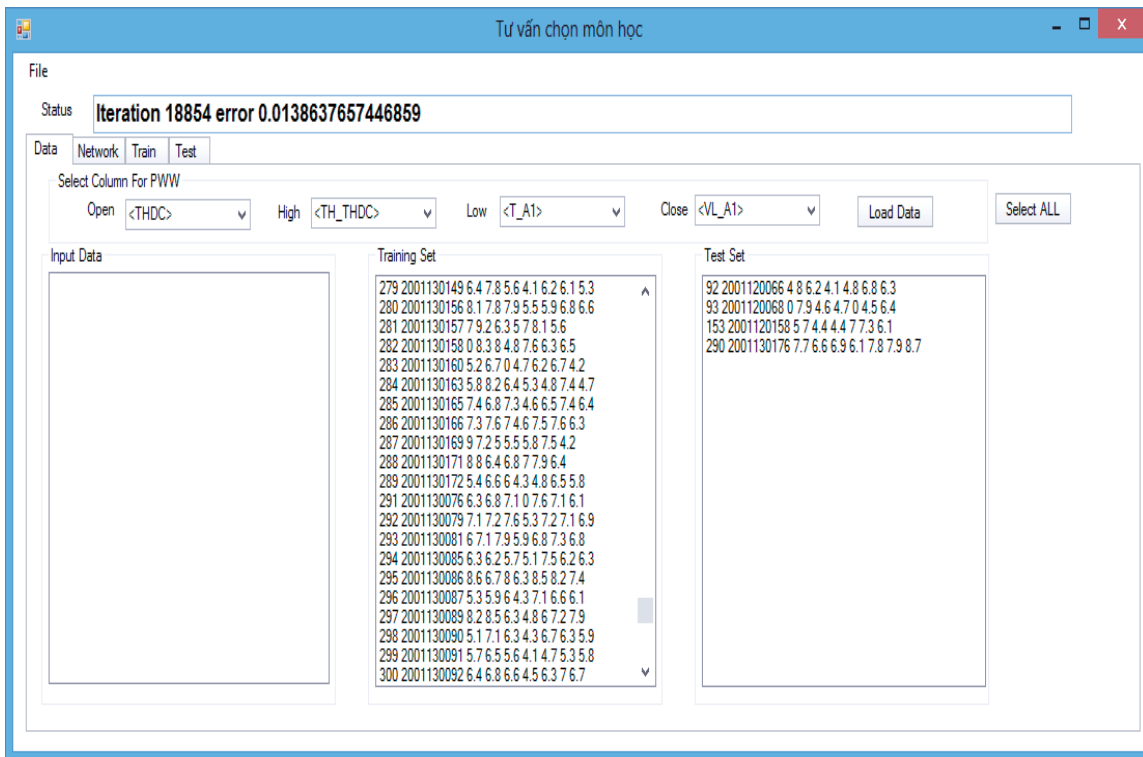
### Quá trình huấn luyện mạng



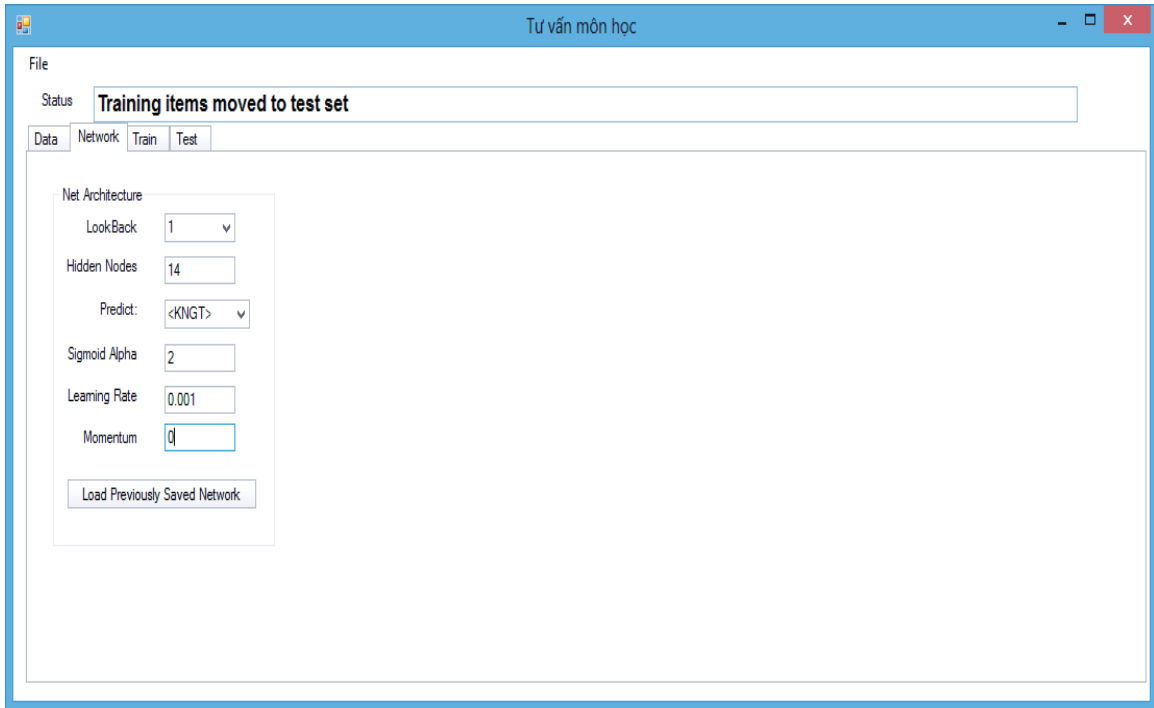
Hình 7. Tải dữ liệu huấn luyện



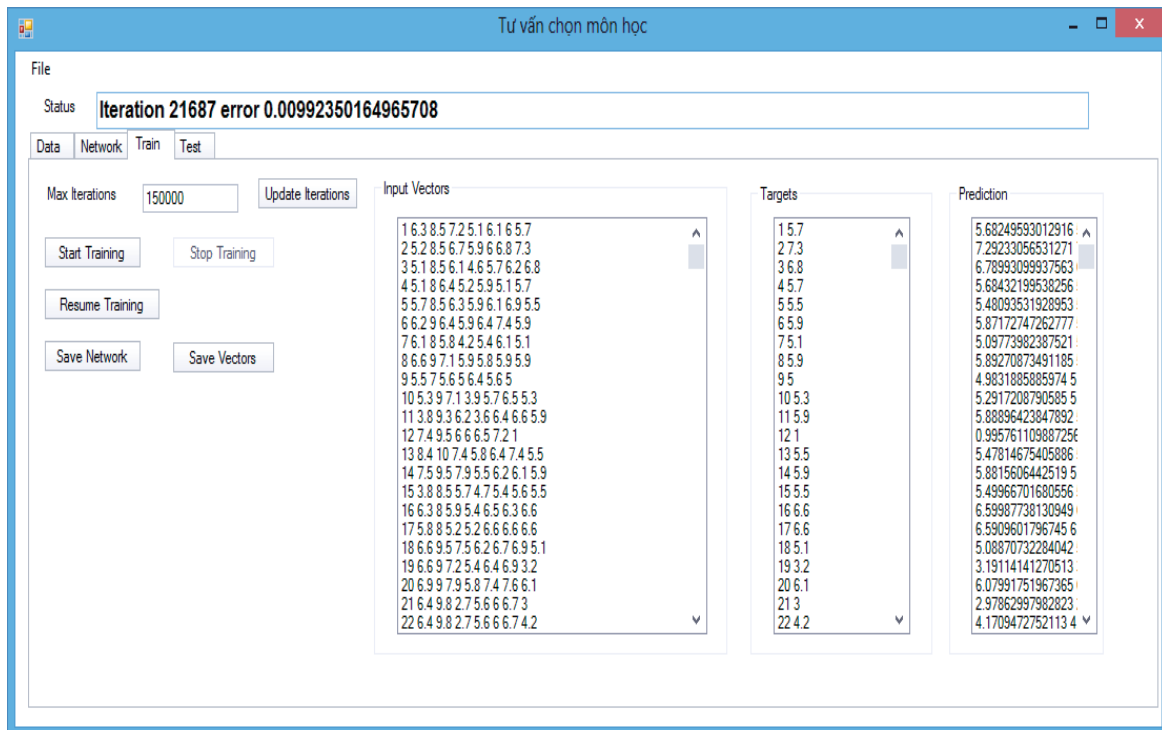
Hình 8. Chọn dữ liệu huấn luyện và dữ liệu kiểm tra



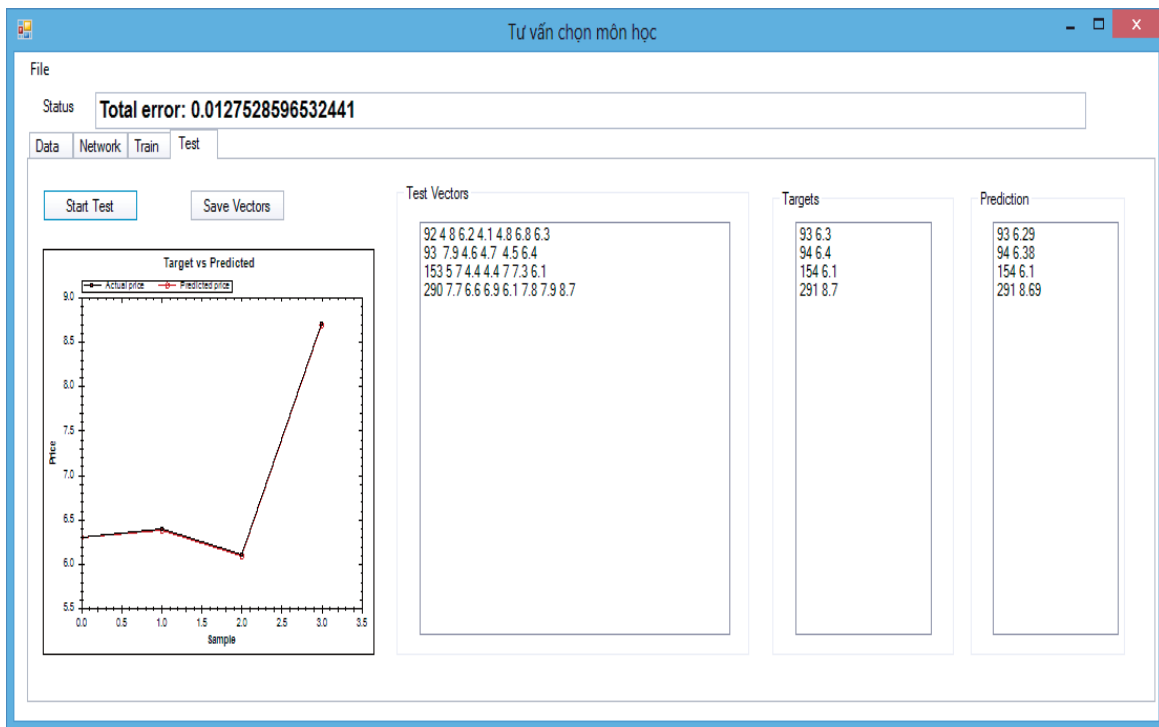
Hình 9. Dữ liệu huấn luyện và kiểm tra



Hình 10. Thông số huấn luyện mạng

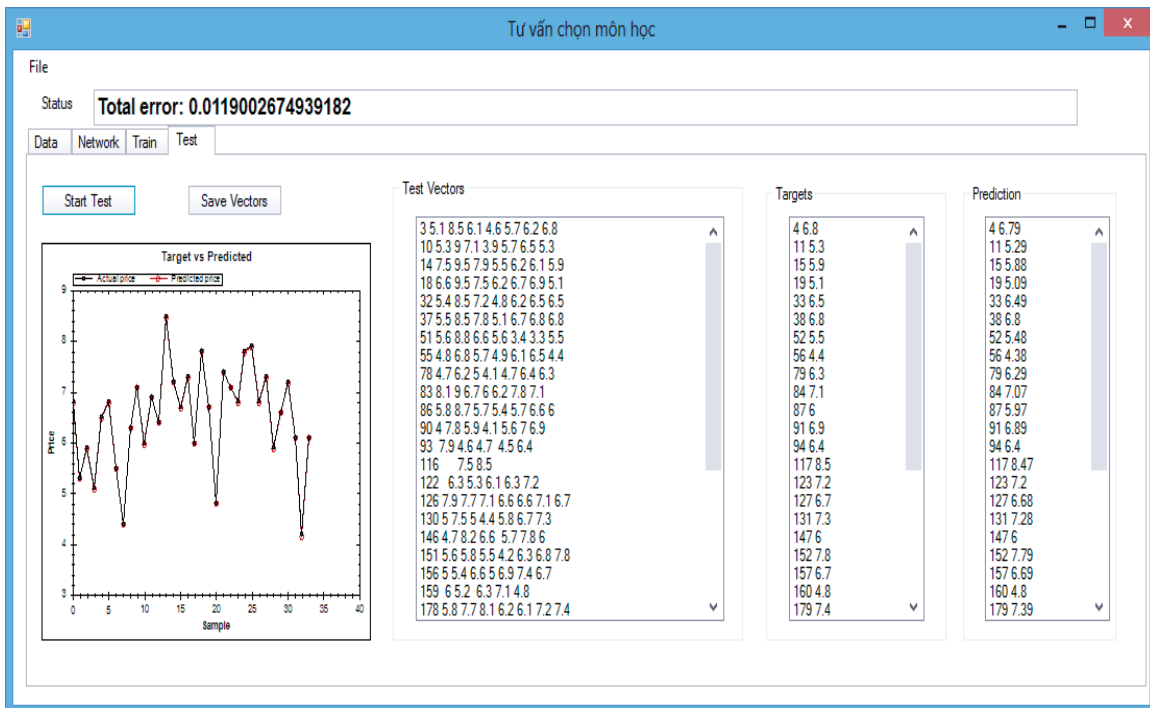


Hình 11. Quá trình huấn luyện



Hình 12. Kết quả kiểm thử

**Nhận xét kết quả thực hiện chương trình**

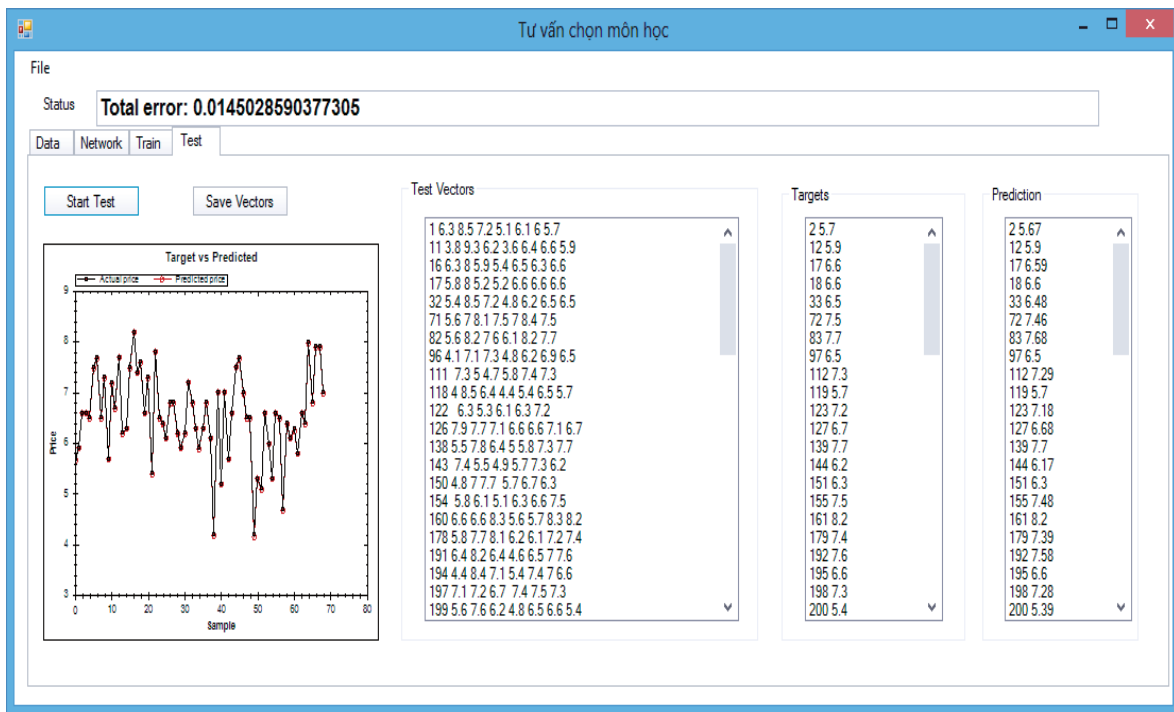


Hình 13. Kết quả thực hiện kiểm thử lần 1



**Bảng 1. So sánh số liệu kiểm thử lần 1**

334 mẫu dữ liệu; Tập huấn luyện 90%; Kiểm thử 10%	
Số lần lặp	20000
Sai số huấn luyện	0.01037
Sai số kiểm thử	0.0119



**Hình 14. Kết quả thực hiện kiểm thử lần 1**

**Bảng 2. So sánh số liệu kiểm thử lần 2**

334 mẫu dữ liệu; Tập huấn luyện 80%; Kiểm thử 20%	
Số lần lặp	20000
Sai số huấn luyện	0.0106
Sai số kiểm thử	0.01450

Với kết quả như trên hoàn toàn chấp nhận được trong quá trình tư vấn điểm môn học cho sinh viên.

## 5. KẾT LUẬN

Việc sử dụng mạng nơron nhân tạo trong tư vấn kết quả học tập sinh viên là một phương pháp hiệu quả, khách quan và khoa học nhằm hỗ trợ cho sinh viên trong quá trình chọn môn học tự chọn. Quá trình cài đặt chương trình tư vấn môn học sử dụng mạng nơron nhân tạo đã thu được một số kết quả nhất định. Kết quả thực nghiệm cho thấy hoàn toàn có khả năng ứng dụng rộng rãi phương pháp này vào thực tế.

### TÀI LIỆU THAM KHẢO

- [1]. Đỗ Phúc, Giáo trình khai thác dữ liệu, Nhà xuất bản đại học quốc gia thành phố Hồ Chí Minh, 2005.
- [2]. Đỗ Thanh Nghị, Giáo trình khai thác dữ liệu, Nhà xuất bản đại học Cần Thơ, 2008.
- [3]. Hà Quang Thụy, Phan Xuân Hiếu, Đoàn Sơn, Nguyễn Trí Thành, Nguyễn Thu Trang, Nguyễn Cẩm Tú, Giáo trình Khai phá dữ liệu Web, NXB Giáo dục, 2009.
- [4]. Bài giảng Khai phá dữ liệu, Đại học Hàng Hải Việt Nam, 2011.
- [5]. Lê Đức Trung, Công nghệ phần mềm, Nhà xuất bản khoa học kỹ thuật, 2005.
- [6]. Pieter Adriaans, Dolf Zantinge, Data Mining, Addison - Wesley, 1998.
- [7]. T. Bhavani, Data Mining: Technologies, Techniques, Tools and Trends. CRC Press 1999.
- [8]. Robert Nisbet, John Elder, Gary Miner, Handbook of Statistical Analysis and Data Mining Applications, Elsevier Inc, 2009.
- [9]. Jiawei Han and Micheline Kamber, Data Mining Concepts and Techniques, Elsevier Inc, 2006.
- [10]. Elmasri, Navathe, Somayajulu, Gupta, Fundamentals of Database Systems (the 4<sup>th</sup> Edition), Pearson Education Inc, 2004.
- [11]. P. N. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison-Wesley, 2006.
- [12]. Ian H., Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2005.
- [13]. M.Kantardzic: Data Mining: Concepts, Models, Method, and Algorithms, John Wiley & Sons, New York, NY, (2003).
- [14]. O'Reilly Media, Inc, 2009, Programming Entity Framework, 1st Edition.
- [15]. Rob Cameron and Dale Michalk, 2004, Building ASP.NET Server Controls.