BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

NGUYỄN ĐĂNG THỊ ÁI MỸ

TRÍCH RÚT TRI THỨC TỪ VĂN BẢN TIẾNG VIỆT TRONG CHẾ BIẾN, BẢO QUẢN RAU QUẢ VÀ ỨNG DỤNG

Chuyên ngành: KHOA HỌC MÁY TÍNH
Mã số: 60.48.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng – Năm 2013

Công trình được hoàn thành tại

ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: PGS.TS. PHAN HUY KHÁNH

Phản biện 1: TS. TRƯƠNG NGỌC CHÂU

Phản biện 2: PGS.TS. TRƯƠNG CÔNG TUẤN

Luận văn đã được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp Thạc sĩ kỹ thuật họp tại Đại học Đà Nẵng vào ngày 23 tháng 02 năm 2013.

Có thể tìm hiểu luận văn tại:

- Trung tâm Thông tin – Học liệu, Đại học Đà Nẵng
- Trung tâm Học liệu, Đại học Bách Khoa

MỞ ĐẦU

1. Lý do chọn đề tài

Ở nước ta, đặc biệt ở các khu công nghiệp, trường học, vấn đề ngộ độc thực phẩm đang ngày càng gia tăng, gây bức xúc trong xã hội. Theo Tổ chức Y tế thế giới (WHO) công bố tại Việt Nam hàng năm có khoảng 8 triệu người bị ngộ độc thực phẩm. Vấn đề ở đây là do người sử dụng thiếu kiến thức trong khâu chế biến và quy trình bảo quản thực phẩm.

Bên cạnh đó nhu cầu sử dụng rau quả hàng ngày rất lớn, gây ra nhiều vấn đề trong quá trình bảo quản và chế biến rau quả như: Nhà nước thiếu quy hoạch trong xu thế đô thị hóa do đó sản xuất rau quả chưa tập trung, phụ thuộc vào điều kiện đất đai. Mặt khác quá trình vận chuyển rau quả phục vụ cho người sử dụng thường xảy ra các vấn đề trong khâu bảo quản, cách chế biến không đúng cũng gây ra ngộ độc thực phẩm...Một phần do kiến thức được đào tạo trong nhà trường chưa thực sự được sử dụng hiệu quả vào thực tế.

Trường Cao đẳng Lương thực – Thực phẩm là trường trọng điểm của Bộ NN&PTNT tại Đà Nẵng. Trường đào tạo đa dạng nhiều ngành nghề bậc Cao Đẳng, Trung cấp. Trong đó ngành nghề đã có truyền thống 35 năm đào tạo là Công nghệ thực phẩm, sinh viên ngành là nguồn nhân lực đáng tin cậy cho các cho các xưởng sản xuất chế biến, xí nghiệp thực phẩm… Dạy tốt, học tốt là tiêu chí hàng đầu của nhà trường và sinh viên. Các môn học: Bảo quản thực phẩm, Công nghệ chế biến rau quả,... rất được SV quan tâm. Nguồn kiến thức về vấn đề này rất đa dạng và phong phú, gây khó khăn cho SV trong việc tìm kiếm được nguồn tri thức đầy đủ và khoa học. Do

đó cần có hệ thống hỗ trợ trích rút tri thức liên quan là một nhu cầu cấp thiết đối với học sinh, sinh viên của nhà trường.

Sự bùng nổ thông tin trên internet hiện nay làm nảy sinh nhu cầu xây dựng các cơ sở tri thức từ nguồn dữ liệu này. Các cơ sở tri thức sẽ cho phép chúng ta quản lý, truy nhập, trao đổi thông tin một cách dễ dàng hiệu quả hơn. Bên cạnh đó, các cơ sở tri thức cũng cho phép máy móc thực hiện những suy diễn trên đó, từ đó tạo ra những tri thức mới phục vụ con người. Để xây dựng các cơ sở tri thức từ khối dữ liệu khổng lồ trên internet hiện nay, vấn đề trích rút ra thông tin (thực thể, quan hệ…) từ các tài liệu là một vấn đề then chốt.

Trong khi đó các hệ thống tìm kiếm hỗ trợ như Google, Yahoo,….vẫn chưa đáp ứng đủ nhu cầu của người sử dụng. Sinh viên phải tự sàng lọc kết quả để thu được tri thức, với vốn kiến thức còn hạn chế không phân biệt được thông tin đúng và sai. Ví dụ: để tìm cách "sản xuất dứa lạnh đông", Google trả về 372.000 kết quả, gây khó khăn cho học sinh trong việc tự sàng lọc và tìm thấy tri thức đúng một cách hiệu quả. Mặt khác, trong quá trình thực hành chế biến và bảo quản rau quả, sinh viên thường gặp nhiều khó khăn trong việc tìm kiếm các nguyên nhân và cách khắc phục lỗi của thành phẩm, hầu hết sinh viên phải tự rà soát lại quá trình chế biến bảo quản để tự tìm ra nguyên nhân và cách khắc phục, làm mất nhiều thời gian và công sức của người học.

2. Mục đích – nhiệm vụ của đề tài

Đề tài "*Trích rút tri thức từ văn bản tiếng Việt trong chế biến, bảo quản rau quả và ứng dụng*" nhằm tìm hiểu về công nghệ tri thức từ đó xây dựng hệ thống hỗ trợ tư vấn cho học sinh sinh viên trong quá trình nghiên cứu học tập và thực hành môn Công nghệ chế biến rau quả tại trường CĐ Lương thực -Thực phẩm Đà Nẵng.

Mục tiêu của đề tài là trích rút tri thức từ các văn bản tiếng Việt trong chế biến bảo quản rau quả tạo ra kho dữ liệu chứa tri thức và dựa vào đó xây dựng hệ chuyên gia để cho phép cập nhật, khai thác dữ liệu một cách dễ dàng, đồng thời tư vấn cách bảo quản và chế biến rau quả hiệu quả nhất.

Nhiệm vụ cụ thể của đề tài :

- Nguyên cứu về công nghệ tri thức.

- Tìm hiểu về tình hình dạy và học môn học Công nghệ chế biến rau quả trong ngành Công nghệ thực phẩm tại trường CĐ Lương thực – Thực phẩm Đà Nẵng.

- Xây dựng cơ sở dữ liệu thành các luật và sự kiện. Tạo thành kho dữ liệu có khả năng cập nhật và truy xuất thông minh dữ liệu có liên quan đến chế biến, bảo quản rau quả.

- Xây dựng hệ thống có giao diện người dùng thân thiện, cho phép người dùng cập nhật dữ liệu thường xuyên để làm giàu kho dữ liệu.

- Kiểm tra đánh giá kết quả thực hiện được.

3. Đối tượng và phạm vi nghiên cứu

3.1. Đối tượng nghiên cứu

Đối tượng nghiên cứu của đề tài bao gồm tìm hiểu về môn học Công nghệ chế biến và bảo quản rau quả tại trường CĐ Lương thực – Thực phẩm ĐN và nhu cầu của học sinh sinh viên trong việc tìm kiếm nguồn kiến thức cho môn học này. Tìm hiểu về tri thức, cách biểu diễn tri thức và hệ chuyên gia.

3.2. Phạm vi nghiên cứu

Phạm vi nghiên cứu của đề tài là phương pháp trích rút tri thức từ văn bản tiếng Việt, xây dựng cơ sở tri thức và xây dựng ứng dụng hỗ trợ tư vấn trong chế biến và bảo quản rau quả phục vụ cho đối

tượng là học sinh sinh viên, giáo viên trong trường CĐ Lương thực – Thực phẩm ĐN. Nghiên cứu cách tạo luật, cơ chế suy diễn, cách biểu diễn và lưu trữ tri thức. Phương thức sử dụng, vận hành và quản lý kho tri thức. Ngôn ngữ lập trình liên quan.

4. Phương pháp nghiên cứu

4.1. Phương pháp tài liệu

Thu thập các tài liệu liên quan đến môn học Công nghệ chế biến rau quả từ giáo viên, thư viện, internet…

Nghiên cứu về tri thức, công nghệ tri thức, cách biểu diễn tri thức để tạo tiền đề xây dựng kho tri thức phục vụ cho hệ thống.

4.2 Phương pháp thực nghiệm

Triển khai xây dựng và cài đặt thử nghiệm, đánh giá kết quả hệ thống hỗ trợ trong bảo quản, chế biến rau quả dựa vào cơ sở tri thức trích rút được.

5. Ý nghĩa khoa học và thực tiễn

Việc thực hiện đề tài giúp cho bản thân hiểu các kiến thức về lĩnh vực công nghệ tri thức, hệ chuyên gia trợ giúp cho học sinh sinh viên trong quá trình nghiên cứu học tập và thực hành trong ngành Công nghệ thực phẩm tại trường CĐ Lương thực – Thực phẩm Đà Nẵng.

6. Bố cục luận văn

Ngoài phần mở đầu, kết luận, tài liệu tham khảo và phụ lục, bố cục luận văn chia làm ba chương được viết tóm lược như sau :

Chương 1. CƠ SỞ LÝ THUYẾT

Trình bày những cơ sở lý thuyết có liên quan đến đề tài : cơ sở tri thức, trích rút tri thức, hệ chuyên gia, kho tri thức.

Chương 2. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG TRÍCH RÚT TRI THỨC TRONG CHẾ BIẾN, BẢO QUẢN RAU QUẢ

Phân tích tình hình dạy và học môn Công nghệ chế biến rau quả tại trường CĐ Lương thực – Thực phẩm Đà Nẵng trong ngành Công nghệ Thực phẩm từ đó đưa ra giải pháp để xây dựng hệ thống.

Chương 3. CÀI ĐẶT VÀ THỬ NGHIỆM HỆ THỐNG

Phân tích chức năng hệ thống, thiết kế kiến trúc hệ thống, xây dựng hệ thống cơ sở dữ liệu (CSDL). Thực hiện xây ứng dụng hệ tư vấn đã thiết kế theo cách thức hệ chuyên gia. Thử nghiệm và đánh giá kết quả chương trình đạt được.

CHƯƠNG 1
CƠ SỞ LÝ THUYẾT

Chương 1 giới thiệu cơ sở lý thuyết liên quan đến đề tài, và là chương làm nền tảng trong việc hỗ trợ xây dựng ứng dụng. Trình bày các khái niệm về tri thức, khai phá tri thức. Một số các ứng dụng cơ sở tri thức để xây dựng thành hệ chuyên gia. Cách thức biểu diễn tri thức trong hệ chuyên gia thông qua các luật và kỹ thuật suy luận trong hệ chuyên gia.

1.1. TỔNG QUAN VỀ TRI THỨC VÀ KHAI PHÁ TRI THỨC

1.1.1. Tri thức

Tri thức hay kiến thức có nhiều ý nghĩa tùy theo văn cảnh, nhưng lúc nào cũng có liên quan với những khái niệm như hiểu biết, ý nghĩa, thông tin, giảng dạy, giáo dục, giao tiếp, diễn tả, học hỏi và kích thích trí óc. Môn học về tri thức được gọi nhận thức luận. Trong nhận thức luận, một định nghĩa phổ biến của tri thức là nó bao gồm ba tiêu chí khả tín, xác thực, và chứng minh được.

Tri thức là:

- Các thông tin, các tài liệu, các cơ sở lý luận, các kỹ năng khác nhau, đạt được bởi một tổ chức hay một cá nhân thông qua các trải nghiệm thực tế hay thông qua sự giáo dục đào tạo; các hiểu biết về lý thuyết hay thực tế về một đối tượng, một vấn đề, có thể lý giải được về nó;

- Là những gì đã biết, đã được hiểu biết trong một lĩnh vực cụ thể hay toàn bộ, trong tổng thể;

- Các cơ sở, các thông tin, tài liệu, các hiểu biết hoặc những thứ tương tự có được bằng kinh nghiệm thực tế hoặc do những tình huống, hoàn cảnh cụ thể. Những tranh cãi về mặt triết học nhìn chung bắt đầu với phát biểu của Plato: tri thức như là những vấn đề đã minh chứng (justified true belief). Tuy nhiên không có một định nghĩa chính xác nào về tri thức hiện nay được mọi người chấp nhận, có thể bao quát được toàn bộ, vẫn còn nhiều học thuyết, các lý luận khác nhau về tri thức.

Tri thức giành được thông qua các quá trình nhận thức phức tạp: quá trình tri giác, quá trình học tập, tiếp thu, quá trình giao tiếp, quá trình tranh luận, quá trình lý luận, hay kết hợp các quá trình này.

1.1.2. Cơ sở tri thức

Cơ sở tri thức chứa các tri thức chuyên sâu về lĩnh vực như chuyên gia. Cơ sở tri thức bao gồm : các sự kiện, các luật, các khái niệm và các quan hệ. Cơ sở tri thức chứa các tri thức để từ đó, máy suy diễn tạo ra câu trả lời cho người sử dụng qua hệ thống giao tiếp.

Hệ cơ sở tri thức là chương trình máy tính được thiết kế để mô hình hóa khả năng giải quyết vấn đề của chuyên gia con người, là hệ thống dựa trên tri thức cho phép mô hình hóa các tri thức của chuyên gia, dùng tri thức để giải quyết vấn đề phức tạp[4].

Ví dụ : kỹ sư tri thức (Knowledge Engineer) là người thiết kế, xây dựng và thử nghiệm hệ chuyên gia

a. Phân loại tri thức

b. Quản trị tri thức

1.1.3. Khai phá tri thức

a. Khai phá tri thức là gì

Kỹ thuật Khai phá tri thức và Khai phá dữ liệu đã và đang được nghiên cứu, ứng dụng trong nhiều lĩnh vực khác nhau ở các nước trên thế giới, tại Việt Nam kỹ thuật này tương đối còn mới mẻ tuy nhiên cũng đang được nghiên cứu và dần đưa vào ứng dụng.

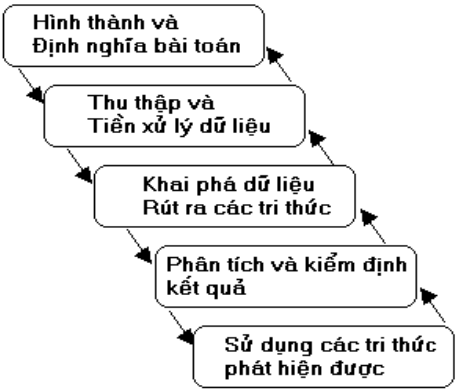
Thông thường chúng ta coi dữ liệu như một dãy các bit, hoặc các số, các ký hiệu, hoặc các "đối tượng" với một ý nghĩa nào đó khi được gửi cho một chương trình dưới một dạng nhất định. Chúng ta sử dụng các bit để đo lường các thông tin và xem nó như là các dữ liệu đã được lọc bỏ các dư thừa, được rút gọn tới mức tối thiểu để đặc trưng một cách cơ bản cho dữ liệu. Chúng ta có thể xem tri thức như là các thông tin tích hợp, bao gồm các sự kiện và các mối quan hệ giữa chúng. Các mối quan hệ này có thể được hiểu ra, có thể được phát hiện, hoặc có thể được học. Nói cách khác, tri thức có thể được coi là dữ liệu có độ trừu tượng và tổ chức cao.

Thông tin và tri thức hiện đang là tiêu điểm của một lĩnh vực mới trong nghiên cứu và ứng dụng về khai phá tri thức (Knowledge Discovery) và khai phá dữ liệu (Data Mining).

Khai phá tri thức trong các cơ sở dữ liệu là một qui trình nhận biết các mẫu hoặc các mô hình trong dữ liệu với các tính năng: hợp thức, mới, khả ích, và có thể hiểu được.

b. Quy trình khai phá tri thức

Qui trình khai phá tri thức được mô tả tóm tắt trên Hình 1:



Hình 1.1 Quy trình khai phá tri thức

1.2. TRÍCH RÚT TRI THỨC TỪ VĂN BẢN TIẾNG VIỆT VÀ KHO TRI THỨC

1.2.1. Trích rút tri thức từ văn bản tiếng Việt

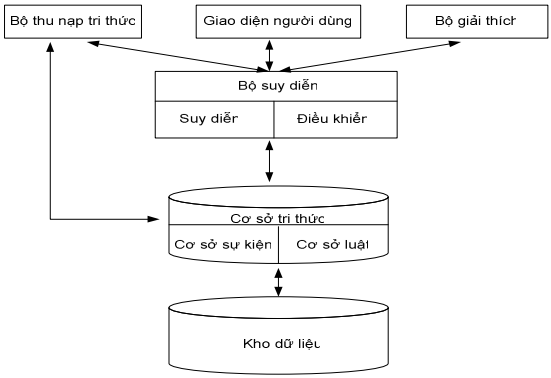
Trích rút tri thức là một bước quan trọng trong quá trình khai phá tri thức. Việc trích rút tri thức quyết định tính hiệu quả, tính khả dụng của hệ thống. Đây chính là bước thứ 3 của quá trình khai phá tri thức đã nói ở trên. Khai phá dữ liệu, trích rút ra tri thức là bước chiếm nhiều thời gian và công sức khi người quản trị thi thức phải thu thập dữ liệu, nghiên cứu lĩnh vực cụ thể liên quan để từ đó trích lọc ra nguồn tri thức khoa học từ các tài liệu thu thập được, hoặc từ chuyên gia trong lĩnh vực đó.

Luận văn chủ yếu tập trung nghiên cứu trích rút tri thức trong văn bản tiếng Việt để xây dựng hệ thống trợ giúp cho học sinh sinh viên trong môn học Công nghệ chế biến rau quả. Do đó chủ yếu thực hiện trích rút tri thức từ văn bản tiếng Việt, thông qua các bài giảng của giáo viên giảng dạy, các đề tài, luận văn liên quan đến chế biến và bảo quản rau quả, các tài liệu tham khảo tiếng Việt trên

internet,… dưới sự giúp đỡ của các giáo viên ngành Công nghệ thực phẩm tại trường CĐ Lương thực – Thực phẩm.

1.2.2. Kho tri thức

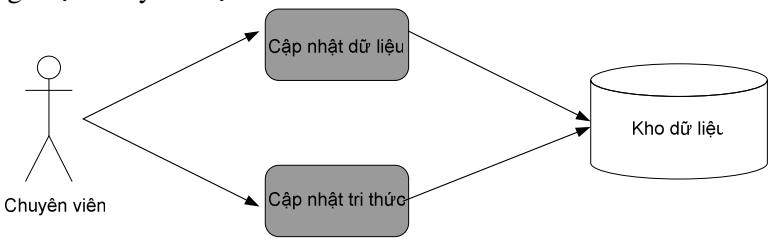
a. Cấu trúc của kho tri thức



Hình 1.3 Mô hình cấu trúc kho tri thức

b. Hệ thống cập nhật, quản lý kho tri thức

Chúng ta cần xây dựng một hệ thống phần mềm để có thể thu nhận, quản lý và đặc tả tri thức. Hệ thống này tạo ra cho người dử dụng có thể dễ dàng tiếp cận kho tri thức. Ngoài ra, hệ thống phải được thiết kế sao cho người sử dụng có thể dễ dàng cập nhật, bổ sung hoặc xử lý dữ liệu.



Hình 1.4 Mô hình hệ thống cập nhật kho dữ liệu

c. Hệ thống quản lý và kho tri thức

Hệ thống này sẽ tạo ra môi trường giao tiếp theo phương thức của hệ chuyên gia. Có hai trường hợp:

Hệ thống sẽ đưa ra các câu hỏi để hướng dẫn người dùng trả lời và dựa vào câu trả lời của người dùng mà hệ thống đưa ra kết quả hợp lý.

Người dùng chọn câu hỏi từ bàn phím , hệ thống sẽ tìm kiếm trả lời tương ứng

1.3. HỆ CHUYÊN GIA

1.3.1 Khái niệm hệ chuyên gia

Hệ chuyên gia (Expert System) là một chương trình máy tính thông minh sử dụng tri thức (knowledge) và các thủ tục suy luận (inference procedure) để giải những bài toán tương đối khó khăn đòi hỏi chuyên gia mới giải được (Theo E.Feigenbaum) [5]

a. Sự phát triển công nghệ hệ chuyên gia

b. Một số lĩnh vực ứng dụng của hệ chuyên gia

c. Kiến trúc tổng quát của hệ chuyên gia

1.3.2. Các đặc trưng và ưu điểm để ứng dụng

a. Các đặc trưng cơ bản của hệ chuyên gia

b. Một số ưu điểm của hệ chuyên gia

1.3.3. Biểu diễn tri thức trong hệ chuyên gia

a. Tri thức được biểu diễn như thế nào

Biểu diễn tri thức bởi các luật sản xuất

Biểu diễn tri thức nhờ mệnh đề logic

Biểu diễn tri thức nhờ mạng ngữ nghĩa

Biểu diễn tri thức nhờ ngôn ngữ nhân tạo

b. Tri thức được biểu diễn thông qua các luật

c. Kỹ thuật suy luận trong hệ chuyên gia

d. Những sai sót thường gặp khi thiết kế hệ chuyên gia

CHƯƠNG 2
PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG TRÍCH RÚT TRI THỨC TRONG CHẾ BIẾN, BẢO QUẢN RAU QUẢ

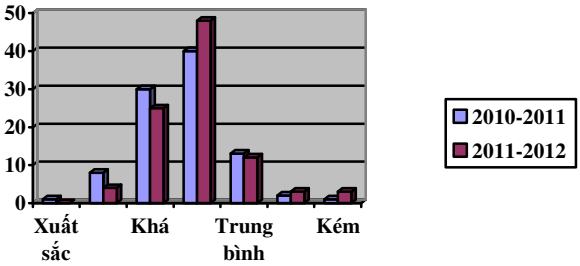
Trong chương 1, luận văn đã trình bày một số khái niệm về công nghệ tri thức, ứng dụng CNTT trên các lĩnh vực và cách xây dựng luật, sự kiện cũng như xây dựng kho tri thức. Chương 2 thực hiện việc phân tích tình hình dạy và học môn chế biến bảo quản rau quả tại trường CĐ Lương thực – Thực phẩm, phát biểu bài toán và trình bày giải pháp trợ giúp cho học sinh sinh viên trong quá trình học tập.

2.1. PHÁT BIỂU BÀI TOÁN

2.1.1. Giới thiệu về trường CĐ Lương thực – Thực phẩm Đà Nẵng

2.1.2. Phân tích tình hình dạy và học môn Công nghệ chế biến rau quả

Thống kê kết quả học tập học sinh khối C22 ngành Kỹ thuật chế biến và bảo quản thực phẩm – bậc TCCN trong 2 năm học vừa qua:



Hình 2.1 Thống kê kết quả học tập ngành Kỹ thuật chế biến và bảo quản thực phẩm

Biểu đồ trên lấy số liệu tổng số học sinh bậc trung cấp chuyên nghiệp ngành Công nghệ thực phẩm khóa 2010-2012, kết quả trên cho thấy rằng kết quả học tập của học sinh của năm thứ 2 thấp hơn so với năm thứ 1, kết quả học tập thấp do ảnh hưởng bởi nhiều yếu tố, một trong các lý do đó là chương trình các môn học chuyên ngành hầu hết tập trung vào năm học thứ 2 của toàn khóa học, cần học sinh hết sức nổ lực để hoàn thành.

2.1.3. Một số khó khăn của học sinh sinh viên trong môn học Chế biến rau quả

Các môn học Công nghệ chế biến rau quả và Thực hành chế biến rau quả ứng dụng trong cuộc sống rất đa dạng và phong phú do nhu cầu sử dụng rau quả của con người hàng ngày rất lớn. Sinh viên trong quá trình học đã tham gia thực hành chế biến nhiều loại rau quả thành thành phẩm với mục đích học tập và sản xuất tại phân xưởng của nhà trường.

Việc sinh viên tham gia sản xuất và nghiên cứu rất được nhà trường quan tâm và khuyến khích, tuy nhiên thành phẩm sau khi sản xuất chỉ tiêu thụ nội bộ, chưa được phổ biến rộng rãi với nhiều nguyên nhân:

- Việc đổi mới phương pháp dạy học mới được đưa vào áp dụng tại trường học sinh còn bỡ ngỡ với phương pháp tự học là chính.

- Học sinh sinh viên của nhà trường đầu vào thấp và phần nhiều thuộc vùng sâu vùng xa, các tỉnh tây nguyên.... nên việc tự tìm hiểu thêm kiến thức trên internet bị hạn chế.

- Sự bùng nổ thông tin trên Internet gây khó khăn cho sinh viên trong việc tự sàng lọc kiến thức trong quá trình tìm kiếm, việc này đòi hỏi rất nhiều thời gian và công sức.

- Một số học sinh ngoài thời gian đi học còn đi làm kiếm tiền chi trả cho cuộc sống hàng ngày nên hạn chế thời gian tự học, khó khăn trong việc tự thu thập tài liệu để trích rút những tri thức có ích cho môn học.

2.2. GIẢI PHÁP TRÍCH RÚT TRI THỨC HỖ TRỢ HỌC SINH SINH VIÊN TRONG CHẾ BIẾN, BẢO QUẢN RAU QUẢ

2.2.1. Giải pháp đề xuất

Bài toán

Qua những phân tích trên ta thấy rằng việc hỗ trợ quản lý dữ liệu trong chế biến và bảo quản rau quả và giúp sinh viên có thể tìm thấy những tri thức cần thiết là nhu cầu thiết yếu. Học sinh sinh viên có thể tra cứu thông tin số liệu về các loại rau quả như: tra cứu về quả dứa (giống dứa, các thành phần vi chất có trong quả dứa, ...), tư vấn cho học sinh sinh viên cách chế biến bảo quản các sản phẩm từ dứa. Để cung cấp các tri thức đúng, khoa học một cách hiệu quả nhất thì cần phải có hệ thống trợ giúp một cách có khoa học.

a. Lý do chọn hệ chuyên gia

b. Mô hình hệ chuyên gia cho bài toán

c. Các bước triển khai xây dựng hệ chuyên gia

Để xây dựng hệ chuyên gia nhằm mục đích hỗ trợ quản lý và tư vấn chọn giống cây trồng theo bài toán. Ta có thể tinh gọn trình tự các bước như sau:

Bước 1 : Xác định mục đích, yêu cầu của người sử dụng

Tại bước này ta cần xác định rõ ràng các câu hỏi. Mục đích sử dụng để làm gi ? Tại sao cần thiết phải xây dựng ? Đối tượng sử dụng là ai ?.

Ngoài ra cần đưa ra dự kiến sử dụng cơ sở tri thức của ứng dụng, giải quyết một số vấn đề công nghệ xử lý tri thức. Ví dụ : trong

bài toán này đối tượng sử dụng là Ban quản lý rừng phòng hộ, họ cần có một hệ thống trợ giúp trong công tác quản lý cũng như tư vấn chọn loại cây trồng và tính dự toán nhằm đưa ra những quyết định trong công tác lập kế hoạch cho dự án hay kế hoạch trồng rừng hàng năm.

Bước 2 : Thu thập tài liệu liên quan đến lĩnh vực xây dựng, cũng như các số liệu thống kê đã có từ nhiều năm trước. Xử lý dữ liệu và xây dựng thành cơ sở tri thức

Tại bước này xác định nguồn tài liệu thu thập. Số liệu được cung cấp từ đâu ? Dữ liệu được lấy từ nguồn nào ? Dữ liệu được trích lọc như thế nào ? Ví dụ: Theo bài toán thì dữ liệu được trích lọc từ các bài giảng môn học Công nghệ chế biến rau quả, các bài báo nông nghiệp, tài liệu tham khảo trên internet, các luận văn…

Bước 3 : Nắm bắt kỹ thuật xây dựng hệ chuyên gia

Ở bước này xây dựng các luật dựa vào nguồn tri thức thu thập được, gồm các bước sau :

- Xác định một số khái niệm quan trọng và mối liên hệ giữa chúng với nhau.
- Xây dựng các tập luật
- Phân tích thiết kế bài toán theo hướng đối tượng
- Xây dựng các lớp và mối quan hệ ràng buộc giữa các lớp.

Mục đích chính ở bước này là phương pháp tiếp nhận tri thức của hệ chuyên gia. Đầu tiên công nghệ tri thức được thu nhận tri thức nhờ đối thoại trực tiếp với con người (chuyên gia), tài liệu trên internet và tài liệu giấy. Sau đó tri thức được biểu diễn (theo luật) tường minh trong cơ sở tri thức, các chuyên gia trao đổi với công nghệ tri thức cho đến khi thỏa mãn nhu cầu bài toán

Bước 4 : Thiết kế giao diện cho hệ chuyên gia

Xây dựng các giao diện sao cho thân thiện với người sử dụng và dễ sử dụng.

Bước 5 : chọn ngôn ngữ để lập trình cho bài toán

Nên chọn ngôn ngữ có nhiều công cụ hỗ trợ, có thể dễ dàng sử dụng những thủ tục có sẵn và có tính tương thích cao. Trong phần này tác giả chọn ngôn ngữ thể hiện C# và ngôn ngữ truy vấn SQL (lưu dữ liệu)

Bước 6 : chạy thử và cải tiến

Cải tiến được tiến hành trong khi mã hóa, nếu phát hiện ra lỗi hoặc có thay đổi mới thì mã cần thay đổi để phù hợp với yêu cầu mới đó.

Hoặc cải tiến khi chạy thử chương trình xảy ra lỗi bên ngoài như việc mở rộng ứng dụng, cần bổ sung thêm luật.

Bước 7: Đánh giá sản phẩm

Sau khi hoàn thành hệ thống được sử dụng trực tiếp bởi người sử dụng từ đó có thể đánh giá và sửa lỗi sai lần cuối trước khi được sử dụng lâu dài và bàn giao sản phẩm.

Tóm lại một hệ chuyên gia có thể mô tả thuật toán như sau

Begin

Chọn bài toán thích hợp

Phát biểu và đặc tả bài toán

If Hệ chuyên gia giải quyết thoả mãn bài toán và có thể sử dụng **Then**

While Bản mẫu chưa được phát triển hoàn thiện **Do**

Begin

Thiết kế bản mẫu

Biểu diễn tri thức

Tiếp nhận tri thức

Phát triển hoàn thiện bản mẫu

End

Hợp thức hoá bản mẫu

Triển khai cài đặt

Hướng dẫn sử dụng

Vận hành

Bảo trì và phát triển

Else

Tìm các tiếp cận khác thích hợp hơn

EndIf

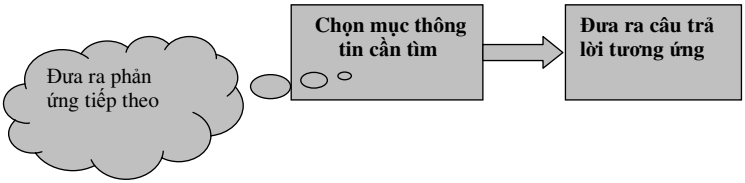
Kết thúc

End

2.2.2. Nhu cầu trợ giúp của học sinh sinh viên

a. Trợ giúp quản lý dữ liệu

Quản lý dữ liệu là nơi có thể tra cứu các thông tin về từng loại rau quả cụ thể theo yêu cầu của học sinh sinh viên, cụ thể là đặc điểm, thành phần hóa học, thành phần dinh dưỡng, cách chế biến, bảo quản, phụ gia trong chế biến của từng loại rau quả tương ứng, cơ cấu công nghiệp chế biến của từng loại rau quả



Ví dụ: học sinh sinh viên chọn nhóm FruitVe, hệ thống yêu cầu người sử dụng nhập vào loại rau quả cần tra cứu, hệ thống đưa

ra dữ liệu về Đặc điểm; Thành phần hóa học; Thành phần dinh dưỡng.

Ví dụ ta nhập vào Dứa, xuất ra Đặc điểm: Tên gọi khác của dứa là thơm, khóm, tên khoa học là *Ananas comosus* là một loại quả nhiệt đới. Quả dứa là trục của bông hoa và các là bắc mọng nước tụ hợp lại. Dứa có các loại: nhóm hoàng hậu (Queen), nhóm Tây Ban Nha (Spanish), nhóm Cayenne. Người sử dụng có thể chọn từng loại giống để xem đặc tính cụ thể, ví dụ: chọn giống dứa Queen thì xuất ra đặc tính: Thịt quả có màu vàng đậm, giòn, mùi thơm tự nhiên, vị ngọt, quả có kích thước nhỏ, khối lượng quả trung bình, mắt quả lồi, ở nước ta gọi là dứa hoa (dứa Tây), là loại dứa có chất lượng cao nhất và hiện đang là nguồn nguyên liệu chính cho chế biến ở Việt Nam.

Thành phần hóa học: độ khô 18%, độ acid 0.51%, chỉ số pH 3.8, đường khử 4,19%, saccharoza 11,59%.

b. Tư vấn cách chế biến và bảo quản

Bài toán tư vấn chế biến và bảo quản rau quả sẽ tư vấn cho người sử dụng cách chế biến, quy trình chế biến và bảo quản phù hợp cho sản phẩm mà người sử dụng yêu cầu. Bài toán liên quan đến các điều kiện sau

Điều kiện có sản phẩm

Hệ thống sẽ cho người sử dụng nhập chọn loại rau quả cần chế biến, nếu trong hệ thống có các cách chế biến về loại rau quả này sẽ tiếp tục cho người sử dụng chọn sản phẩm cần chế biến từ loại rau quả đó, nếu không sẽ trả về thông báo.

Để chọn sản phẩm cần chế biến, thể hiện như sau:

Chọn sản phẩm cần chế biến

Người sử dụng nhập chọn sản phẩm cần chế biến từ loại rau quả tương ứng, nếu không tìm thấy loại sản phẩm thì hiển thị thông báo đồng thời gợi ý cho người sử dụng chọn các sản phẩm chế biến khác của cùng loại rau quả mà người sử dụng đã chọn từ ban đầu.

c. Các chức năng cơ bản

Chương trình được thiết kế thành 3 chức năng cơ bản sau :

1. Chức năng quản trị hệ thống
2. Chức năng Quản lý dữ liệu
3. Chức năng tư vấn cách chế biến và bảo quản
4. Chức năng tư vấn nguyên nhân và cách khắc phục hư hỏng sản phẩm trong quá trình chế biến và bảo quản.

2.3. MÔ TẢ ỨNG DỤNG VÀ MÔ HÌNH DỮ LIỆU HOẠT ĐỘNG

2.3.1. Mô tả ứng dụng

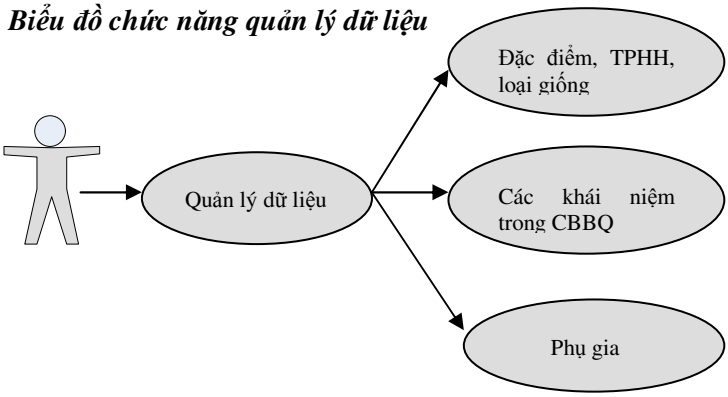
a. Mô tả các lớp

b. Xây dựng thuộc tính và cá thể

2.3.2. Thiết kế mô hình dữ liệu của hệ tư vấn

2.3.3. Xây dựng qui trình hoạt động hệ thống

a. Biểu đồ chức năng quản lý dữ liệu



Hình 2.8 Biểu đồ ca sử dụng chức năng quản lý dữ liệu

b. Biểu đồ chức năng tư vấn chế biến và bảo quản

c. Biểu đồ chức năng tư vấn tìm nguyên nhân và cách khắc phục lỗi

2.4. XÂY DỰNG CÁC LUẬT VÀ CÂU TRUY VẤN

2.4.1. Xây dựng hệ thống các luật và sự kiện

a. Tập luật chế biến và bảo quản

Chọn nguyên liệu

1. **Nếu** chọn nguyên liệu cho sản xuất dứa khoanh nước đường, **thì**:

Giống dứa Queen

Quả to, có hình trụ

Có độ chín từ một hàng mắt đến nữa quả

2. **Nếu** chọn nguyên liệu cho sản xuất dứa lạnh đông, **thì**:

Quả tươi tốt, không mốc meo, bầm dập

Có độ chín từ 3 mắt đến 2/3 quả

b. Tập luật tư vấn tìm nguyên nhân và biện pháp khắc phục lỗi sản phẩm

Trước khi tư vấn, người sử dụng cần nhập tên của sản phẩm rau quả cần kiểm tra vào hệ thống.

Xây dựng luật dựa trên sản phẩm "dứa khoanh nước đường"

Chương trình sẽ cho phép người quản trị bổ sung luật, điều kiện và định mức nếu luật phát sinh có điều kiện mới.

Sau đây là một số luật cơ bản trong hệ thống :

Luật 1 : Nếu tình trạng hộp sản phẩm dứa khoanh nước đường:

Vi sinh vật gây bệnh và hở ghép mí

➔Nguyên nhân : Lỗi xảy ra trong quá trình ghép mí

TimNguyenNhan(X,XX) ^

coTinhTrang("Vi sinh vật gây bệnh và hở ghép mí")

→TimNguyenNhan("Lỗi xảy ra trong quá trình ghép mí")

Luật 2: Nếu nguyên nhân Lỗi xảy ra trong quá trình ghép mí

→Cách khắc phục: Kiểm tra áp suất, nhiệt độ và thời gian ghép mí theo ngưỡng tới hạn. Dùng thiết bị để hiệu chỉnh, sửa chữa. Kiểm tra độ hở mí của tất cả các hộp được ghép mí trong thời gian vi phạm, hủy bỏ các hộp bị hở mí.

Luật 3 : Nếu tình trạng hộp sản phẩm dứa khoanh nước đường:

Vi sinh vật gây bệnh và ghép mí kín.

→Nguyên nhân: Lỗi xảy ra trong quá trình hấp thanh trùng.

TimNguyenNhan(X,XX) ^

coTinhTrang("Vi sinh vật gây bệnh và ghép mí kín")

→TimNguyenNhan("Lỗi xảy ra trong quá trình hấp thanh trùng")

2.4.2. Áp dụng các tập luật vào trong hệ thống chương trình

Trong chương trình chúng ta áp dụng hệ thống luật. Tập luật được đánh số thứ tự, mỗi luật có một tên luật và nhận mã luật làm khóa. Mỗi luật được xác định một tập điều kiện, cho phép người quản trị nhập bổ sung chúng từ bàn phím.

Ví dụ: luật 1 có điều kiện là vi sinh vật gây bệnh và hở ghép mí thì trong cơ sở dữ liệu có tên *luật 1*, tên điều kiện *tình trạng lỗi,* đối chiếu để lấy ra nguyên nhân và biện pháp khắc phục.

2.4.3. Thiết kế một số tư vấn (tạo ra tiền đề và kết quả)

Công việc thiết kế tư vấn do người quản trị tạo ra, thông tin để tạo ra tư vấn được lấy từ nhiều nguồn khác nhau, có thể lấy từ internet hay từ những tài liệu khác nhau. Sau đó trích lọc và tạo thành luật.

CHƯƠNG 3
CÀI ĐẶT THỬ NGHIỆM HỆ THỐNG

Trong chương trước, luận văn đã trình bày bài toán cần tư vấn và đưa ra giải pháp ứng dụng hệ chuyên gia bằng ngôn ngữ truy vấn SQL được thiết kế trên Microsoft viusal studio để trợ giúp học sinh sinh viên trong môn học Công nghệ chế biến rau quả dựa trên hệ thống trích rút tri thức. Chương này tập trung vào việc thiết kế, xây dựng, thử nghiệm ứng dụng rồi từ đó đánh giá kết quả đạt được của luận văn.

3.1. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1.1. Các mô hình chức năng hệ thống

a. Chức năng quản lý dữ liệu

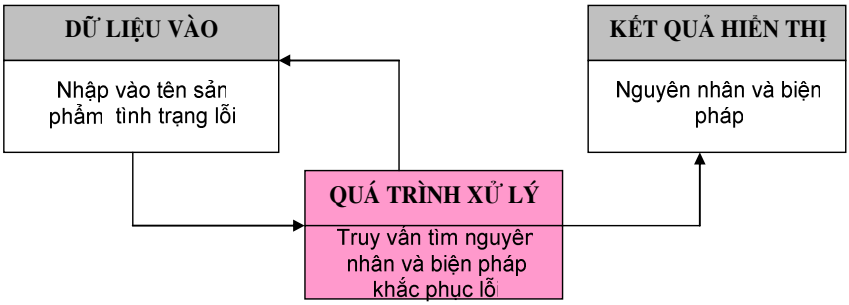
Chức năng này cho phép người sử dụng tra cứu thông tin về các loại rau quả, bao gồm đặc điểm, đặc tính, thành phần hóa học, các khái niệm trong chế biến và bảo quản.

b. Chức năng tư vấn cách chế biến và bảo quản rau quả

Chức năng này người sử dụng nhập vào loại rau quả cần tư vấn cách chế biến, nếu tồn tại trong hệ thống thì sẽ nhập tiếp chọn mục sản phẩm cần chế biến và bảo quản, ngược lại hệ thống sẽ hiển thị thông báo.

c. Chức năng tư vấn tìm nguyên nhân và cách khắc phục lỗi sản phẩm

Chức năng có mục đích là tư vấn cho người sử dụng tìm ra nguyên nhân và biện pháp khắc phục lỗi sản phẩm . Chức năng được thể hiện qua mô hình sau:



Hình 3.3 Mô hình chức năng tư vấn nguyên nhân và biện pháp khắc phục lỗi

Đầu vào:

Nhập vào tên sản phẩm cần tư vấn, hệ thống sẽ kiểm tra xem sản phẩm đó đã tồn tại trong hệ thống chưa, nếu có hệ thống yêu cầu nhập chọn tình trạng lỗi sản phẩm, các tình trạng lỗi sản phẩm được gợi ý trong danh sách tình trạng lỗi có thể xảy ra đối với sản phẩm tương ứng.

Đầu ra :

Hệ thống suy diễn trả về kết quả là nguyên nhân gây ra tình trạng lỗi đó và các biện pháp phòng ngừa.

Dòng sự kiện:

Mỗi một luật chứa trong bảng dbo.bangluat chứa các sự kiện maluat, matinhtrang của một luật.

Ví dụ: Luật 1 : Nếu tình trạng hộp sản phẩm dứa khoanh nước đường: Vi sinh vật gây bệnh và hở ghép mí

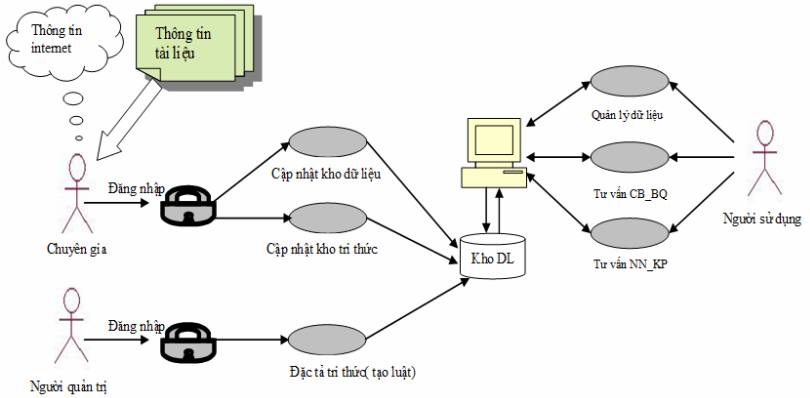
➔Nguyên nhân : Lỗi xảy ra trong quá trình ghép mí

Duyệt danh sách tình trạng lỗi, sau đó chọn tình trạng lỗi xảy ra với sản phẩm được chọn, tương ứng với mã tình trạng, đối chiếu vào dbo.bangluat ta được mã luật và xuất ra màn hình nguyên nhân và cách khắc phục.

d. Chức năng cập nhật kho dữ liệu

e. Chức năng cập nhật cơ sở tri thức

3.1.2. Mô hình kiến trúc tổng thể hệ thống



Hình 3.4 Mô hình kiến trúc tổng thể hệ thống

3.2. XÂY DỰNG HỆ THỐNG ỨNG DỤNG

3.2.1. Mô-đun 1 Quản lý dữ liệu

3.2.2. Mô-đun 2 Tư vấn chế biến và bảo quản rau quả

3.2.3. Mô-đun 3 Tư vấn tìm nguyên nhân và khắc phục lỗi sản phẩm

3.2.4. Mô-đun 4 Quản trị

3.3. MÔI TRƯỜNG CÔNG CỤ CÀI ĐẶT

3.4. ĐÁNH GIÁ KẾT QUẢ

3.4.1. Đánh giá kết quả chương trình

3.4.2. Kịch bản vận dụng hệ thống

KẾT LUẬN

1. Kết luận

Qua thời gian nghiên cứu và học hỏi kinh nghiệp từ những người đã nghiên cứu trước đó về lĩnh vực này, tôi đã hoàn thành luận văn và có thể tóm tắt như sau :

Luận văn nêu lên được bài toán và giải pháp để giải quyết bài toán. Xây dựng được hệ thống tư vấn cách chế biến và bảo quản rau quả dựa vào các tri thức trích rút được từ các bài giảng, luận văn.. giải quyết được những khó khăn cơ bản của học sinh sinh viên trong quá trình học tập môn học Công nghệ chế biến rau quả tại trường Cao đẳng Lương thực – Thực phẩm.

Về mặt lý thuyết, đã nêu được giải pháp ứng dụng công nghệ hệ chuyên gia để xây dựng tri thức thành luật, sử dụng ngôn ngữ SQL để lưu trữ kho dữ liệu và chương trình visual studion 2010 để trình bày bài toán, trình bày được một cách tổng quát các khái niệm về tri thức, trích rút tri thức, quản lý tri thức, hệ chuyên gia,…

Về mặt thực tiễn, đề tài đã đáp ứng được các mục tiêu đề ra, tạo ra kho dữ liệu tri thức trợ giúp học sinh sinh viên trong quá trình học tập.

2. Hướng phát triển đề tài

Bên cạnh những vấn đề đạt được nêu ở trên, đề tài cần được cần được phát triển thêm một số vấn đề sau :

Bổ sung thêm thông tin (hạng mục) cần quản lý một cách chi tiết, cụ thể hơn

Cập nhật bổ sung nhiều luật để người sử dụng có thể được tư vấn chế biến và bảo quản nhiều loại sản phẩm rau quả hơn, học sinh sinh viên có thể được tư vấn khắc phục được tất cả các lỗi sản phẩm.

Tạo thêm chức năng hướng dẫn sử dụng và đưa đề tài thực sự đi vào ứng dụng.