# TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP.HCM

ỨNG DỤNG LÝ THUYẾT BAYES TRONG PHÂN LỚP ĐỂ XÂY DỰNG HỆ THỐNG HỖ TRỢ

CHẨN ĐOÁN BỆNH VIÊM PHẾ QUẢN

# MÔN HỌC: KHAI THÁC DỮ LIỆU GIÁO VIÊN HƯỚNG DẪN: NGÔ DƯƠNG HÀ

**THÀNH VIÊN NHÓM: 1. LÝ HOÀNG PHI DŨNG - 2001181065 2. PHAN THỊ NGỌC DUNG - 2001181062**

**TP.HCM - Năm 2020**

# MỞ ĐẦU

1. **Lí do chọn đề tài**

Hàng năm trên thế giới có hơn 4 triệu người tử vong do các bệnh liên quan đến đường hô hấp, trong đó phổ biến nhất là viêm phế quản (Bronchitis). Viêm phế quản là bệnh thường gặp ở trẻ em, bất kể lứa tuổi nào và thời tiết nào, đặc biệt là trẻ ở thành thị cũng như ở các nơi tập trung dân cư đông đúc thì tỉ lệ bệnh còn cao hơn. Thế giới hiện có khoảng 300 triệu người mắc viêm phế quản và khoảng 250.000 trường hợp tử vong vì viêm phế quản mỗi năm. Tỷ lệ mắc viêm phế quản ước tính khoảng 6-8% ở người lớn và khoảng 10-12% ở trẻ em dưới 15 tuổi.

Chính vì vậy, hiểu biết một số nguyên nhân và cách đề phòng của căn bệnh này sẽ làm giảm thiểu số ca bệnh, đề phòng được các biến chứng và nâng cao chất lượng cuộc sống. Đồng thời cũng làm giảm nguy cơ dẫn đến các biến chứng do bệnh này gây ra như hen phế quản, ung thư phế quản, viêm phổi, áp xe phổi, suy hô hấp cấp….

## **Mục tiêu đối tượng của đề tài**

Từ nhu cầu thực tiễn trong cuộc sống về khám chữa bệnh cộng với những động lực đã nêu trên, đề tài mong muốn ứng dụng công nghệ thông tin vào lĩnh vực y khoa và đặc biệt là chuyên khoa hô hấp để chẩn đoán bệnh viêm phế quản. Đề tài cũng sẽ tìm hiểu và nghiên cứu về các mô hình biểu diễn tri thức và những phương pháp, kĩ thuật xây dựng hệ chuyên gia để vận dụng vào việc xây dựng hệ hỗ trợ chẩn đoán bệnh một cách tự động. Như vậy, đề tài hướng tới việc xây dựng một hệ hỗ trợ chẩn đoán bệnh viêm phế quản có cơ sở tri thức từ các bác sĩ chuyên khoa Nội - Hô hấp để hỗ trợ cho các bác sĩ chưa có nhiều kinh nghiệm hoặc các bác sĩ không chuyên về Nội - Hô hấp có thể tham khảo kinh nghiệm chẩn đoán và điều trị của các bác sĩ chuyên khoa nhằm đưa ra những quyết định chính xác trong trường hợp chưa có bác sĩ chuyên khoa kịp thời.

Chương trình hỗ trợ cho các bác sĩ không thuộc chuyên khoa Nội – Hô hấp, các bác sĩ đa khoa tuyến huyện chưa có kiến thức chuyên sâu về bệnh viêm phế quản. Chương trình hỗ trợ bác sĩ trong việc kết luận bệnh.

## **Ý nghĩa khoa học và thực tiễn**

Vấn đề chẩn đoán và điều trị bệnh viêm phế quản đã được các bác sĩ chuyên khoa áp dụng thường xuyên trong việc khám bệnh thực tế tại bệnh viện. Tuy nhiên việc đưa tri thức của các chuyên gia này vào mô hình biểu diễn tri thức để ứng dụng nó trong tin học thì chưa được tiến hành, vì vậy việc ứng dụng công nghệ tri thức để biểu diễn và ứng dụng vào lĩnh vực y khoa để xây dựng một hệ thống chẩn đoán tự động có ý nghĩa thực tiễn rất lớn.

Mục Lục

[MỞ ĐẦU 1](#_Toc59625315)

[**1.** **Lí do chọn đề tài** 1](#_Toc59625316)

[**2.** **Mục tiêu đối tượng của đề tài** 1](#_Toc59625317)

[**3.** **Ý nghĩa khoa học và thực tiễn** 2](#_Toc59625318)

[Chương 1. Giới Thiệu Về Thuật Toán Bayes Và Hệ Thống 4](#_Toc59625320)

[**1.** **Phương pháp phân lớp Naïve Bayes** 4](#_Toc59625321)

[1.1. Định lý Bayes 4](#_Toc59625322)

[**1.2.** Thuật toán Naïve Bayes 4](#_Toc59625323)

[**2.** **Giới thiệu hệ thống** 4](#_Toc59625324)

[**2.1.** Mục đích, đối tượng sử dụng 4](#_Toc59625325)

[**2.2.** Chức năng của hệ thống 5](#_Toc59625328)

[2.3. Triệu chứng của viêm phế quản 6](#_Toc59625331)

[*2.3.1.* *Viêm phế quản cấp* 6](#_Toc59625332)

[Chương 2. CÀI ĐẶT VÀ KIỂM THỬ 7](#_Toc59625336)

[**1.** **Các thành phần của hệ thống** 7](#_Toc59625337)

[*1.1.* *Form chính của hệ thống* 7](#_Toc59625338)

[*1.2.* *Form Chuẩn Đoán Bệnh và Bảng Huấn Luyện* 8](#_Toc59625339)

[KẾT LUẬN VÀ KIẾN NGHỊ 9](#_Toc59625340)

[**1.** **Kết quả đạt được** 9](#_Toc59625341)

[**2.** **Một số hạn chế** 9](#_Toc59625342)

[**3.** **Hướng phát triển** 9](#_Toc59625343)

Chương 1. Giới Thiệu Về Thuật Toán Bayes Và Hệ Thống

## **Phương pháp phân lớp Naïve Bayes**

* 1. Định lý Bayes

Tính xác suất xảy ra của một sự kiện ngẫu nhiên A khi biết sự kiện liên quan B đã xảy ra.

* + - * Xác suất này được kí hiệu là P(A\B)
      * Đọc là “*xác suất của A nếu có B*”

### Thuật toán Naïve Bayes

* Mô hình phân lớp Naïve Bayes
* Thuật toán Naïve Bayes
* Các bước thực hiện thuật toán phân lớp Naïve Bayes
  + - 1. Bước 1: Huấn luyện Naïve Bayes (dựa vào tập dữ liệu), tính P(Ci) và P(Xk|Ci).
      2. Bước 2: Phân lớp Xnew= (x1, x2, …, xn). Xnew ta cần tính xác suất thuộc từng phân lớp khi đã biết trước Xnew. Xnew được gán vào lớp có xác suất lớn nhất theo công thức: 𝑚𝑎𝑥𝐶𝑖∈𝐶(𝑃(𝐶𝑖)∐ 𝑃 𝑛 𝑘=1 (𝑥𝑘|𝐶𝑖))

1. **Giới thiệu hệ thống**

### Mục đích, đối tượng sử dụng

* + - * Ứng dụng những thành quả của công nghệ thông tin vào lĩnh vực y tế, góp phần phục vụ tốt việc tìm hiểu nghiên cứu đánh giá chất lượng y tế tại các phòng khám bệnh.
      * Hỗ trợ công tác khám và chữa bệnh của bệnh nhân và bác sĩ hoàn toàn chính xác, đáng tin cậy.
      * Triển khai ứng dụng thực tiễn tại các phòng khám bệnh, trạm y tế…
* Dữ liệu vào ra của hệ chẩn đoán như sau:

### Dữ liệu đầu vào:

Một số triệu chứng lâm sàng và cận lâm sàng.

### Output:

Đưa ra cho người dùng kết luận bệnh hoặc là đợi xác nhận của chuyên gia và có thể xem chi tiết tỷ lệ xác suất bệnh mắc phải từ những triệu chứng đầu vào dựa trên thuật toán Bayes.

### Chức năng của hệ thống

#### Yêu cầu chức năng

Gồm các chức năng: chẩn đoán bệnh, đăng nhập, cập nhật thông tin các luật, cập nhật tên bệnh, cập nhật triệu chứng, lưu trữ luật, lưu trữ triệu chứng bệnh, lưu trữ tên các bệnh, lưu trữ thông tin người sử dụng.

#### Yêu cầu phi chức năng

* + - * + Tính tiện dụng: Giao diện thân thiện dễ sử dụng.
        + Tính hiệu quả: Đảm bảo khả năng truy xuất nhanh đến các màn hình và khả năng kiểm soát lỗi tốt.
        + Tính tương thích: Hỗ trợ các hệ điều hành Window
        + Hỗ trợ nhiều người dùng truy cập cùng lúc.
        + Mã hóa các thông tin cần được bảo mật.
        + Thiết kế ứng dụng theo mô hình 3 lớp.

### Triệu chứng của viêm phế quản

* + 1. *Viêm phế quản cấp*

##### Triệu chứng lâm sàng

* + Giai đoạn ướt: Cảm giác rát bỏng sau xương ức giảm và hết, khó thở nhẹ, có thể còn sốt ho. Nghe phổi có ran rít, ran ngáy, ran ẩm to và vừa hạt. Diễn biến 4 – 10 ngày thì khỏi hẳn. Ho nhiều đờm, khạc đờm nhầy, hoặc đờm mủ xanh hoặc vàng (khi bội nhiễm). Có trường hợp ho khan kéo dài vài tuần. Có thể bệnh bắt đầu rầm rộ biểu hiện sốt cao, ho nhiều, ho ra máu… Nghe phổi có ran ẩm.

##### Triệu chứng cận lâm sàng

* + X- Quang phổi: rốn phổi đậm.
  + Xét nghiệm máu: số lượng bạch cầu tăng, tốc độ máu lắng

tăng.

* + 1. *Viêm phế quản mạn*

##### Triệu chứng lâm sàng

**Bảng 2.1. Phân biệt khí phế thủng và viêm phế quản mạn**

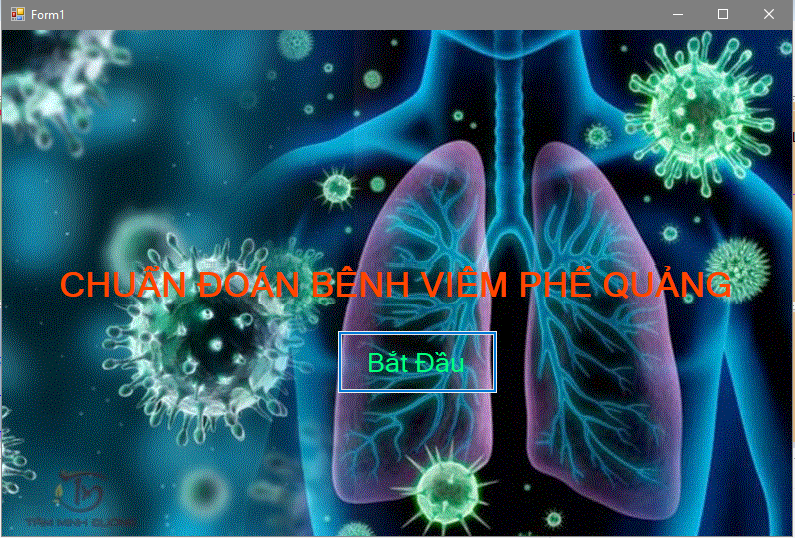
|  |  |  |
| --- | --- | --- |
|  | **Khí phế thủng** | **Viêm phế quản mạn** |
| Khó thở | Nặng | Vừa |
| Ho | Có sau khó thở | Có trước khó thở |
| Viêm đường thở | Ít | Thường xuyên |
| Suy hô hấp | Giai đoạn cuối | Từng đợt cấp |
| X- Quang | Giãn phổi, tăng sáng | Hình ảnh "Phổi bẩn" |
| Sức cản đường thở  (Raw) | Tăng nhẹ | Tăng nhiều |

# Chương 2. CÀI ĐẶT VÀ KIỂM THỬ

## **Các thành phần của hệ thống**

### *Form chính của hệ thống*

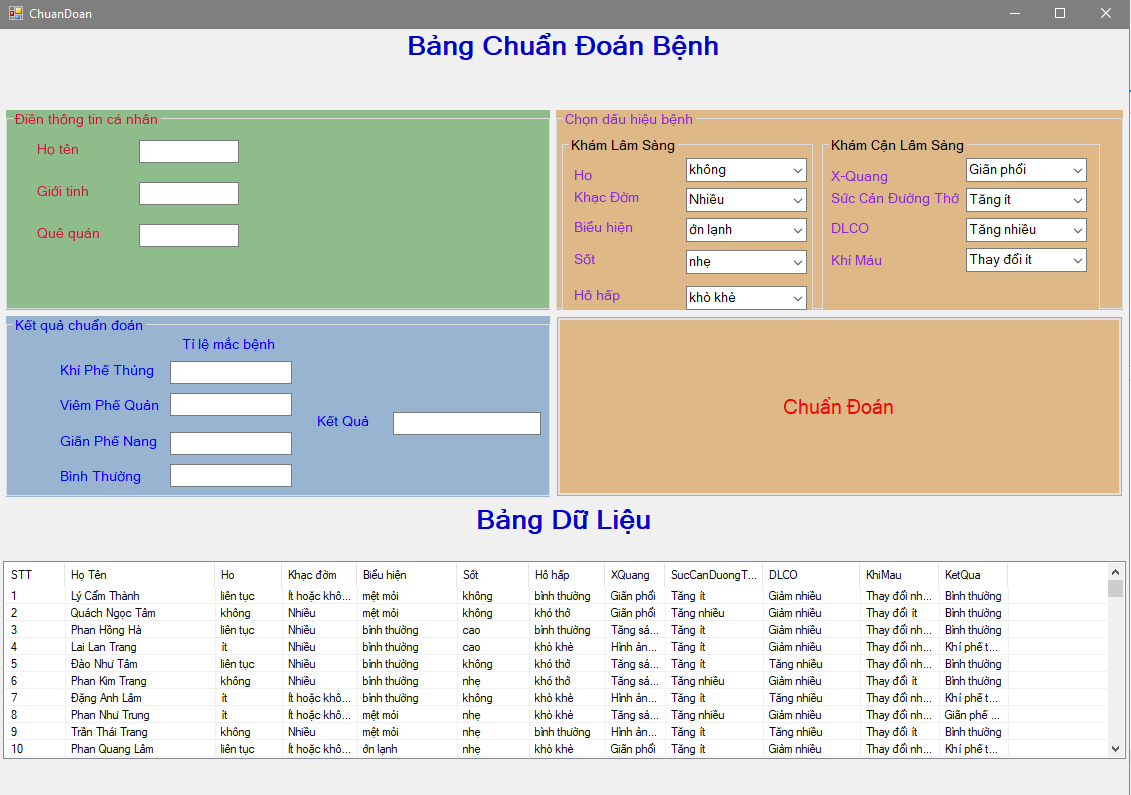
Giao diện chính của hệ thống với các chức năng cho người dùng đăng nhập vào hệ thống hoặc là chỉ có thể xem phần hỗ trợ chẩn đoán



Hình 1. Màn hình Form chính của hệ thống

### *Form Chuẩn Đoán Bệnh và Bảng Huấn Luyện*

Màn hình cho phép người sử dụng và chuyên gia có thể xem hỗ trợ chẩn đoán bệnh. Người dùng có thể chọn các triệu chứng bệnh sau đó nhấn vào nút chẩn đoán để xem kết quả.



Hình 2. Màn hình chuẩn đoán của hệ thống

# KẾT LUẬN VÀ KIẾN NGHỊ

1. **Kết quả đạt được**

Kết quả của khóa luận tốt nghiệp là xây dựng một hệ hỗ trợ chẩn đoán bệnh viêm phế quản dựa trên lý thuyết Bayes phù hợp với thực tế khám bệnh tại bệnh viện và dữ liệu sử dụng là dữ liệu thực tế được thu thập tại phòng khám đa khoa trường ĐH Trà Vinh. Hệ thống này đã đáp ứng giúp ích cho các bác sĩ trẻ mới ra trường và các bác sĩ không chuyên khoa có thêm kinh nghiệm trong việc chẩn đoán bệnh viêm phế quản cũng như học hỏi được kinh nghiệm chẩn đoán của các bác sĩ chuyên khoa lâu năm.

Hệ thống hoạt động dựa vào máy suy diễn và tập luật đã được thu thập từ tri thức chuyên gia để đưa ra kết quả chẩn đoán. Tri thức của chuyên gia đã trãi qua nhiều năm kinh nghiệm nghiên cứu về bệnh viêm phế quản. Tất cả các tri thức thu thập được là hoàn toàn chính xác, đáng tin cậy và đang được sử dụng trong thực tế tại phòng khám.

## **Một số hạn chế**

* Trong thế giới thực, hầu như bất khả thi khi các triệu chứng của người bệnh không nằm trong luật thì lúc này hệ thống không tìm ra được bệnh.
* Không có độ ưu tiên cho từng triệu chứng, vì trong thực tế bác sĩ chỉ cần một số triệu chứng quan trọng thì có thể kết luận được bệnh.

## **Hướng phát triển**

* Cần có sự ưu tiên cho từng triệu chứng.
* Phát triển thành dạng Web phổ biến cho cộng đồng.
* Có thêm chức năng quản lý thông tin bệnh nhân và bệnh án để có thể chẩn đoán thêm chính xác hơn.