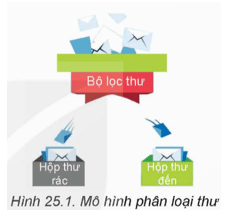
# Bài 25: Làm quen với Học máy

**Giải Tin học 12 Bài 25: Làm quen với Học máy**  
**Khởi động trang 134 Tin học 12**: Khi truy cập tài khoản thư điện tử, ngoài các thư trong Hộp thư đến (Inbox) em có thể thấy nhiều thư được tự động phân loại vào Hộp thư rác (Spam). Hãy quan sát Hình 25.1 và cho biết việc phân loại này được thực hiện như thế nào.  
  
**Lời giải:**  
Việc phân loại thư vào Hộp thư rác (Spam) thông qua quá trình sau:  
- Hộp thư đến (Inbox): Đây là nơi nhận và hiển thị tất cả các thư đến.  
- Bộ lọc thư: Bộ lọc thư là một công cụ hoặc hệ thống được sử dụng để xác định xem một thư điện tử có phải là thư rác hay không. Bộ lọc này có thể sử dụng các quy tắc và thuật toán để phân tích nội dung và thông tin của thư.  
- Thư rác: Các thư được xác định là thư rác sẽ được tự động chuyển vào Hộp thư rác (Spam). Điều này giúp người dùng tách biệt giữa các thư quan trọng và thư không mong muốn hoặc có tính spam.  
  
**Hoạt động 1 trang 134 Tin học 12**: Bộ lọc thư điện tử, thường là tập hợp các quy tắc, được thiết kế để phát hiện và đánh dấu các thư rác trước khi chúng được chuyển vào hộp thư của người dùng. Có quy tắc chỉ đơn giản là trong nội dung hoặc tiêu đề thư có các cụm từ đáng ngờ như “miễn phí”, “giảm giá”, “rẻ bất ngờ".... hay địa chỉ thư của người gửi hoặc địa chỉ của máy chủ gửi thư thuộc vào một “danh sách đen” xác định. Theo em, có thể xây dựng các bộ lọc thư này bằng cách nào?  
**Lời giải:**  
Để xây dựng bộ lọc thư điện tử, có thể sử dụng các phương pháp học máy.  
**Câu hỏi 1 trang 136 Tin học 12**: Chọn phương án đúng. Học máy là:  
A. Chương trình máy tính có khả năng đưa ra quyết định hay dự đoán dựa trên dữ liệu.  
B. Khả năng phân tích dữ liệu thu nhận được của máy tính để đưa ra dự đoán hoặc quyết định dựa trên các quy tắc được xác định rõ ràng.  
C. Việc sử dụng các phương pháp và kĩ thuật cho phép máy tính học từ dữ liệu để đưa ra dự đoán hoặc quyết định mà không cần lập trình cụ thể.  
D. Chương trình máy tính có khả năng tự cải thiện hiệu suất thực hiện nhiệm vụ thông qua việc cập nhật các dữ liệu mới sau khi hoàn thành nhiệm vụ đó nhiều lần.  
**Lời giải:**  
Đáp án C là định nghĩa chính xác và toàn diện nhất về học máy vì:  
- Sử dụng các phương pháp và kỹ thuật: Học máy không chỉ là một phương pháp đơn giản mà nó bao gồm một loạt các kỹ thuật và phương pháp, bao gồm học có giám sát, học không giám sát, và học bán giám sát.  
- Máy tính học từ dữ liệu: Trong học máy, máy tính không chỉ đơn giản là thực hiện các quy tắc được xác định trước (như trong lựa chọn B), mà nó thực sự học từ dữ liệu. Điều này có nghĩa là máy tính sẽ phân tích dữ liệu và điều chỉnh các tham số của mô hình để tối ưu hóa hiệu suất của nó.  
- Đưa ra dự đoán hoặc quyết định: Mục tiêu cuối cùng của học máy là để máy tính có khả năng đưa ra dự đoán hoặc quyết định dựa trên dữ liệu mà nó đã học được. Điều này có thể bao gồm việc dự đoán giá cổ phiếu, phân loại email là spam hay không, hoặc thậm chí là tự lái xe.  
- Không cần lập trình cụ thể: Trong học máy, không cần phải lập trình cụ thể các quy tắc hoặc logic cho máy tính như trong lựa chọn A. Thay vào đó, máy tính tự học từ dữ liệu thông qua các thuật toán và kỹ thuật học máy.  
  
**Câu hỏi 2 trang 136 Tin học 12**: Tại sao cần chia dữ liệu học máy thành hai phần: dữ liệu huấn luyện và dữ liệu kiểm tra?  
**Lời giải:**  
Dữ liệu huấn luyện và dữ liệu kiểm tra được chia ra để đánh giá hiệu suất của mô hình học máy:  
- Dữ liệu huấn luyện được sử dụng để huấn luyện mô hình, tức là mô hình học từ dữ liệu này để tìm ra các mẫu và quy luật trong dữ liệu.   
- Dữ liệu kiểm tra được sử dụng để đánh giá hiệu suất của mô hình đã được huấn luyện.  
  
**Hoạt động 2 trang 136 Tin học 12**: Trong Mục 1 em đã được giới thiệu một số ứng dụng thực tế của Học máy. Hãy kể tên một vài công việc mà ngày nay không thể thiếu vai trò của Học máy.  
**Lời giải:**  
Một số công việc mà ngày nay không thể thiếu vai trò của Học máy:  
- Lọc thư rác.  
- Chẩn đoán bệnh.  
- Phân tích thị trường.  
- Nhận dạng tiếng nói.  
- Nhận dạng chữ viết.  
- Dịch tự động.  
**Câu hỏi trang 139 Tin học 12**: Vai trò quan trọng của Học máy trong các lĩnh vực khác nhau được thể hiện như thế nào?  
**Lời giải:**  
Học máy có vai trò quan trọng trong nhiều lĩnh vực khác nhau và thể hiện sự quan trọng của nó qua các khía cạnh sau:  
- Xử lý dữ liệu lớn: Học máy giúp xử lý lượng lớn dữ liệu trong thời gian thực một cách nhanh chóng và hiệu quả. Nó có khả năng tìm ra các mẫu và xu hướng quan trọng trong dữ liệu, giúp nhận diện thông tin hữu ích và tri thức ẩn trong dữ liệu đó.  
- Tự động hóa nhiệm vụ phức tạp: Học máy cho phép tự động hóa các nhiệm vụ phức tạp mà trước đây đòi hỏi sự can thiệp của con người. Ví dụ, trong lĩnh vực lọc thư rác, học máy giúp xây dựng mô hình phân loại thư rác một cách tự động, giảm công sức và thời gian của người dùng.  
- Chẩn đoán và dự đoán: Trong lĩnh vực y tế, học máy được sử dụng để xây dựng các mô hình chẩn đoán bệnh và dự đoán tình trạng sức khoẻ. Điều này giúp cải thiện khả năng chuẩn đoán của bác sĩ và đưa ra phương án điều trị phù hợp cho bệnh nhân.  
- Phân tích và dự báo thị trường: Học máy có thể phân tích dữ liệu thị trường từ nhiều nguồn khác nhau để xác định xu hướng, dự báo biến động giá cả và hỗ trợ quyết định kinh doanh. Điều này giúp người đầu tư và nhà kinh doanh hiểu rõ hơn về thị trường và đưa ra các chiến lược đầu tư dựa trên thông tin phân tích.  
- Nhận dạng và xử lý âm thanh và hình ảnh: Học máy giúp nhận dạng và xử lý các dạng bức ảnh, tiếng nói và chữ viết tay. Ví dụ, trong việc nhận dạng tiếng nói, học máy giúp xây dựng các mô hình nhận dạng và phân biệt tiếng nói của các người nói khác nhau.  
- Dịch tự động: Học máy được sử dụng để xây dựng các mô hình dịch tự động, giúp dịch văn bản và tiếng nói từ một ngôn ngữ sang ngôn ngữ khác. Điều này giúp loại bỏ rào cản ngôn ngữ và tạo điều kiện cho giao tiếp và trao đổi thông tin đa ngôn ngữ.  
  
**Luyện tập trang 139 Tin học 12**: Tại sao có thể nói Học máy có vai trò không thể thiếu trong các ứng dụng mà dữ liệu không ngừng thay đổi theo thời gian? Hãy chỉ ra một vài minh họa cụ thể.  
**Lời giải:**  
- Bởi vì Học máy là một công cụ hữu ích, là một phần quan trọng của trí tuệ nhân tạo, đóng góp vào sự phát triển của nhiều ứng dụng thực tế.  
- Một vài ví dụ cụ thể:  
+ Xử lý dữ liệu lớn và thời gian thực: Học máy cho phép xử lý một lượng lớn dữ liệu một cách nhanh chóng và hiệu quả. Chương trình học máy có thể tự động xác định các mẫu và xu hướng quan trọng trong dữ liệu, giúp chúng ta hiểu rõ hơn về sự biến đổi và thay đổi trong dữ liệu theo thời gian.  
+ Tự động hóa nhiệm vụ phức tạp: Trước đây, nhiều nhiệm vụ phức tạp đòi hỏi sự can thiệp của con người. Nhưng với học máy, chúng ta có thể tự động hoá các nhiệm vụ này. Ví dụ, học máy có thể tự động phát hiện thư rác trong email hoặc phân loại ảnh dựa trên nội dung.  
+ Bổ sung tri thức và hỗ trợ chuyên gia: Học máy giúp chuyên gia và nhà nghiên cứu xây dựng và bổ sung tri thức. Chúng ta có thể học từ dữ liệu và áp dụng tri thức này vào các lĩnh vực như y học, tài chính, và khoa học tự nhiên.  
+ Tự động cập nhật mô hình: Học máy không ngừng bổ sung dữ liệu và tự động cập nhật mô hình đã được huấn luyện. Điều này đặc biệt quan trọng trong các ứng dụng có dữ liệu đa dạng và thay đổi liên tục theo thời gian, ví dụ như phát hiện các mẫu thư rác mới, triệu chứng bệnh mới, hay bản dịch ngôn ngữ mới.  
  
**Vận dụng trang 139 Tin học 12**: Ngoài hai phương pháp học máy cơ bản nêu trong bài, một số tài liệu còn đề cập tới phương pháp học bán giám sát và học tăng cường. Hãy tìm hiểu về các phương pháp học máy này trên Internet.  
**Lời giải:**  
- Học bán giám sát (Supervised Learning):  
+ Trong học bán giám sát, mô hình được đào tạo trên dữ liệu có nhãn (gần nhãn) và không có nhãn (không gắn nhãn).  
+ Dữ liệu gần nhân là những ví dụ mà chúng ta biết kết quả mong muốn (ví dụ: hình ảnh của mèo hoặc chó đã được gắn nhãn là “mèo" hoặc "chó").  
+ Mục tiêu là học cách dự đoán đúng kết quả cho các ví dụ mới chưa gặp trước đó.  
+ Học bán giám sát thường được sử dụng trong xử lý ngôn ngữ tự nhiên, phân loại hình ảnh, và nhiều ứng dụng khác.  
- Học tăng cường (Reinforcement Learning):  
+ Trong học tăng cường, tác nhân (agent) tương tác với môi trường và nhận phần thưởng (hoặc hình phạt) dựa trên hành động của nó.  
+ Mục tiêu là học cách thực hiện các hành động để tối đa hóa phần thưởng trong môi trường.  
+ Học tăng cường thường được sử dụng trong chế tạo robot, chơi trò chơi, và các tác vụ tối ưu hóa.