

LỜI CAM ĐOAN

Tôi xin cam đoan: Luận văn thạc sỹ chuyên ngành Khoa học máy tính, tên đề tài “Nghiên cứu hỗ trợ chuẩn đoán bệnh lao dựa vào học máy” là công trình nghiên cứu, tìm hiểu và trình bày do tôi thực hiện dưới sự hướng dẫn khoa học của **PGS.TS. Đỗ Năng Toàn**, Viện Công nghệ Thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Kết quả tìm hiểu, nghiên cứu trong luận văn là hoàn toàn trung thực, không vi phạm bất cứ điều gì trong luật sở hữu trí tuệ và pháp luật Việt Nam. Nếu sai, tôi hoàn toàn chịu trách nhiệm trước pháp luật.

Tất cả các tài liệu, bài báo, khóa luận, công cụ phần mềm của các tác giả khác được sử dụng lại trong luận văn này đều được chỉ dẫn tường minh về tác giả và đều có trong danh mục tài liệu tham khảo.

Thái Nguyên, ngày 30 tháng 6 năm 2022.

Tác giả luận văn
Nguyễn Hữu Khánh

LỜI CẢM ƠN

Tác giả xin chân thành cảm ơn **PGS.TS. Đỗ Năng Toàn**, Viện Công nghệ Thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, là giáo viên hướng dẫn khoa học đã hướng dẫn tác giả hoàn thành luận văn này, xin được cảm ơn các thầy, cô giáo trường Đại học công nghệ thông tin và truyền thông nơi tác giả theo học và hoàn thành chương trình cao học đã nhiệt tình giảng dạy và giúp đỡ.

Xin cảm ơn Trung tâm Đào tạo Từ xa - Đại học Thái Nguyên nơi tác giả công tác đã tạo mọi điều kiện thuận lợi để tác giả có thời gian, tâm trí để hoàn thành nhiệm vụ nghiên cứu và chương trình học tập.

Và cuối cùng xin cảm ơn gia đình, bạn bè, đồng nghiệp đã động viên, giúp đỡ tác giả trong suốt thời gian học tập, nghiên cứu và hoàn thành luận văn này.

Xin chân thành cảm ơn.

Thái Nguyên, ngày 30 tháng 06 năm 2022

Tác giả luận văn
Nguyễn Hữu Khánh

Mục lục

Lời cam đoan.....	i
Lời cảm ơn	ii
Mở đầu.....	1

MỞ ĐẦU

Ngày nay, khi xã hội ngày càng phát triển, việc đưa máy tính vào sử dụng, phục vụ cho công việc đời sống của con người đã sản sinh ra một khối lượng dữ liệu lớn và phức tạp (big data), được số hóa và lưu trữ trên máy tính. Những tập dữ liệu lớn này có thể bao gồm các dữ liệu có cấu trúc, không có cấu trúc và bán cấu trúc. Đó có thể là dữ liệu thông tin bán hàng trực tuyến, lưu lượng truy cập trang web, thông tin cá nhân, thói quen hoạt động thường ngày của con người.v.v. Chúng chứa đựng nhiều thông tin quý báu mà khi khai thác hợp lý sẽ trở thành tri thức, tài sản mang lại giá trị lớn. Thách thức đặt ra cho con người là phải đưa ra các phương pháp, thuật toán và công cụ hợp lý làm sao để phân tích được lượng dữ liệu lớn như vậy.

Người ta nhận thấy máy tính có khả năng phân tích, xử lý khối dữ liệu lớn và phức tạp, tìm ra các mẫu và quy luật, vượt quá khả năng, tốc độ tính toán ghi nhớ của bộ não con người. Khái niệm học máy từ đó hình thành. Ý tưởng cơ bản của học máy là máy tính có thể học hỏi, học tự động theo kinh nghiệm [?]. Máy tính phân tích lượng lớn dữ liệu, tìm thấy các mẫu, quy tắc ẩn trong dữ liệu, sử dụng các quy tắc đó để mô tả dữ liệu mới một cách tự động và liên tục cải thiện.

Học máy có rất nhiều ứng dụng, bao gồm nhiều lĩnh vực. Máy tìm kiếm sử dụng học máy để xây dựng mối quan hệ tốt hơn giữa các cụm từ tìm kiếm và các trang web. Bằng cách phân tích nội dung của các trang web, công cụ tìm kiếm có thể xác định từ nào là cụm từ quan trọng nhất trong việc xác định một trang web nhất định và họ có thể sử dụng cụm từ này để trả thông tin kết quả phù hợp cho cụm từ tìm kiếm nhất định [?]. Công nghệ nhận dạng

hình ảnh cũng sử dụng học máy để xác định các đối tượng cụ thể, chẳng hạn như khuôn mặt [?]. Đầu tiên thuật toán học máy phân tích hình ảnh có chứa một đối tượng nhất định. Nếu được cung cấp đủ hình ảnh cho quá trình này, thuật toán có thể xác định được hình ảnh có chứa đối tượng đó hay không [?]. Ngoài ra học máy có thể được sử dụng để hiểu loại sản phẩm mà khách hàng quan tâm, bằng cách phân tích các sản phẩm trong quá khứ mà người dùng đã mua. Máy tính có thể đưa ra đề xuất các sản phẩm khách hàng có thể mua với xác suất cao [?]. Tất cả những ví dụ trên đều có nguyên tắc cơ bản giống nhau: Máy tính xử lý và học cách xác định dữ liệu, sau đó sử dụng kiến thức này để đưa ra quyết định về dữ liệu trong tương lai.

Mặc dù đã có cố gắng nỗ lực, song luận văn không tránh khỏi những thiếu sót do năng lực và thời gian hạn chế. Em chân thành mong muốn lắng nghe những đóng góp, góp ý của thầy, cô, bạn bè, đồng nghiệp để luận văn được cải thiện tốt hơn.

Em xin chân thành cảm ơn.