

THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://www.youtube.com/watch?v=yVR4auBC5LU>
- Link slides (dạng .pdf đặt trên Github):
https://github.com/nguyenminhhieuuit/CS2205.APR2023/blob/main/nhan_dang_tuoi_va_gioi_tinh_dua_tren_mo_hinh_hoc_sau.pdf
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*

<ul style="list-style-type: none">● Họ và Tên: Nguyễn Minh Hiếu● MSSV: 230201010 	<ul style="list-style-type: none">● Lớp: CS2205.APR2023● Tự đánh giá (điểm tổng kết môn): 9/10● Số buổi vắng: 0● Số câu hỏi QT cá nhân: 10● Link Github: https://github.com/mynameuit/CS2205.APR2023/
--	--

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA) NHẬN DẠNG TUỔI VÀ GIỚI TÍNH DỰA TRÊN MÔ HÌNH HỌC SÂU
TÊN ĐỀ TÀI TIẾNG ANH (IN HOA) AGE AND GENDER IDENTIFICATION BASED ON DEEP LEARNING MODEL
TÓM TẮT <i>(Tối đa 400 từ)</i> <p>Trong thời đại mà công nghệ trí tuệ nhân tạo đang bùng nổ, các ứng dụng của học sâu đã trở thành công cụ thiết yếu trong nhiều lĩnh vực, đặc biệt là trong nhận diện hình ảnh và phân tích dữ liệu. Nghiên cứu này giải quyết bài toán dự đoán độ tuổi và giới tính từ hình ảnh của ứng viên, nhằm hỗ trợ quy trình tuyển dụng trong các doanh nghiệp. Vấn đề đặt ra là việc ứng viên khai man thông tin cá nhân như độ tuổi và giới tính, gây khó khăn cho nhà tuyển dụng trong việc xác minh và lựa chọn ứng viên phù hợp. Bằng cách áp dụng phương pháp học sâu, cụ thể là mạng nơ-ron tích chập (CNN), nghiên cứu đã xây dựng một mô hình có khả năng phân tích và dự đoán thông tin cá nhân từ hình ảnh ứng viên. Kết quả thử nghiệm cho thấy mô hình đạt độ chính xác cao trong việc dự đoán độ tuổi và giới tính, từ đó phát triển một hệ thống hỗ trợ tuyển dụng hiệu quả. Hệ thống này giúp tiết kiệm thời gian, chi phí và nâng cao độ tin cậy trong quy trình tuyển dụng.</p>
GIỚI THIỆU <i>(Tối đa 1 trang A4)</i> <p>Sự phát triển kinh tế - xã hội phụ thuộc vào nhiều yếu tố khác nhau, nhưng tập trung chủ yếu vào nguồn nhân lực. Dù có máy móc, thiết bị, nhưng chúng trở nên vô dụng nếu thiếu sự can thiệp của con người. Nguồn nhân lực đóng vai trò trung tâm và quan trọng nhất trong phát triển kinh tế - xã hội của đất nước, đặc biệt trong giai đoạn đẩy mạnh công nghiệp hóa, hiện đại hóa và hội nhập kinh tế quốc tế.</p> <p>Tuyển dụng tốt sẽ tạo đội ngũ nhân lực phù hợp với công việc và vị trí, từ đó nâng cao</p>

hiệu quả công việc và chất lượng nguồn nhân lực của doanh nghiệp. Đồng thời giúp tiết kiệm được chi phí đào tạo nhân lực, tránh tình trạng ứng viên khai gian dối độ tuổi lao động gây ra vi phạm luật lao động hiện hành dẫn tới phải liên tục tuyển dụng nhân sự mới, doanh nghiệp lại phải tốn chi phí và thời gian để tuyển dụng lại nhân sự mới.

Bằng cách áp dụng các kỹ thuật của trí tuệ nhân tạo kết hợp với mạng nơ-ron tích chập (CNN), chúng tôi đã áp dụng các giải pháp nhằm xây dựng hệ thống hỗ trợ công tác tuyển dụng thông qua dự đoán độ tuổi và giới tính, từ đó đưa ra nhận định ứng viên có phù hợp theo độ tuổi lao động hiện hành giúp nhà tuyển dụng tránh được sự khai man của các ứng viên.

Input: Hình ảnh khuôn mặt của ứng viên được trích xuất từ CV hồ sơ ứng tuyển



Hình 1. Hình ảnh minh họa đầu vào 1 (dưới độ tuổi lao động)

Output: Thông tin dự đoán về độ tuổi và giới tính của ứng viên.



Thông tin ứng viên tự khai

Họ tên: Nguyễn Thị Lan

Tuổi: 18 tuổi

Giới tính: Nữ

Kết quả dự đoán

Giới tính: Nữ

Dự đoán ứng viên có độ tuổi từ 0 đến 14 tuổi không phù hợp độ tuổi lao động.
Ứng viên nằm trong vòng nghi vấn.

Làm mới

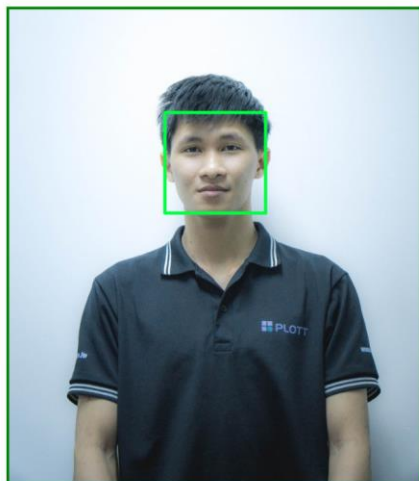
Hình 2. Kết quả nghi vấn dưới độ tuổi lao động

Input: Hình ảnh khuôn mặt của ứng viên được trích xuất từ CV hồ sơ ứng tuyển



Hình 3. Hình ảnh minh họa đầu vào 2(trong độ tuổi lao động)

Đầu ra (Output): Dự đoán giới tính và độ tuổi.



Thông tin ứng viên tự khai

Họ tên: Nguyễn Minh Hiếu

Tuổi: 25 tuổi

Giới tính: Nam

Kết quả dự đoán

Giới tính: Nam

Dự đoán ứng viên có độ tuổi từ 15 đến 61 tuổi phù hợp với độ tuổi lao động.

Làm mới

Hình 4. Kết quả trong độ tuổi lao động

Input: Hình ảnh khuôn mặt của ứng viên được trích xuất từ CV hồ sơ ứng tuyển



Hình 5. Hình ảnh minh họa đầu vào 3(trên độ tuổi lao động)

Đầu ra (Output): Dự đoán giới tính và độ tuổi.



Thông tin ứng viên tự khai

Họ tên: La Văn Hà

Tuổi: 58 tuổi

Giới tính: Nam

Kết quả dự đoán

Giới tính: Nam

Dự đoán ứng viên có độ tuổi cao, lớn hơn 61 tuổi không phù hợp độ tuổi lao động.
Ứng viên nằm trong vòng nghi vấn.

Làm mới

Hình 6. Kết quả nghi vấn trên độ tuổi lao động

Kết quả của quá trình này đó là đề án đã xây dựng được một ứng dụng hỗ trợ công tác tuyển dụng. Ứng dụng đã được cài đặt, thử nghiệm tại phòng nhân sự Công ty Long Vân System Solution giúp cho quá trình tuyển dụng trở nên hiệu quả hơn.

MỤC TIÊU

(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)

Xây dựng mô hình học sâu: Phát triển một mô hình mạng nơ-ron tích chập (CNN) có khả năng dự đoán chính xác độ tuổi và giới tính từ hình ảnh.

Phát triển hệ thống hỗ trợ tuyển dụng: Tạo ra một ứng dụng web tích hợp mô hình dự đoán, giúp nhà tuyển dụng xác định thông tin cá nhân của ứng viên nhanh chóng và chính xác.

Đánh giá hiệu quả: Kiểm chứng và đánh giá hiệu quả của mô hình và hệ thống trên các bộ dữ liệu công khai và dữ liệu thực tế, đảm bảo tính khả thi và độ chính xác.

NỘI DUNG VÀ PHƯƠNG PHÁP

(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)

Nội dung

1. Tổng quan về phương pháp học sâu và ứng dụng của nó trong dự đoán độ tuổi và giới tính.
2. Phân tích các phương pháp xử lý hình ảnh và thị giác máy tính liên quan.
3. Xây dựng và huấn luyện mô hình CNN dựa trên bộ dữ liệu Adience.

Phương pháp

Thu thập và tiền xử lý dữ liệu: Sử dụng bộ dữ liệu Adience, tiến hành các bước tiền xử lý như loại bỏ các đặc trưng không cần thiết, chỉnh kích thước ảnh.

Xây dựng mô hình: Áp dụng kiến trúc mạng AlexNet, xây dựng mô hình CNN với các lớp tích chập và lớp kết nối đầy đủ.

Huấn luyện và đánh giá mô hình: Sử dụng các kỹ thuật tối ưu hóa và hàm mất mát phù hợp, huấn luyện mô hình trên tập dữ liệu huấn luyện và kiểm tra hiệu quả trên tập dữ liệu kiểm tra.

Phát triển hệ thống: Xây dựng ứng dụng web hỗ trợ tuyển dụng, tích hợp mô hình dự đoán vào hệ thống, cho phép nhập thông tin và hình ảnh ứng viên, thực hiện dự

đoán và hiển thị kết quả.

Trước đây đã có nhiều nghiên cứu và giải pháp được đề xuất cho việc dự đoán tuổi và giới tính. Các nghiên cứu trước đây thường sử dụng các phương pháp truyền thống, yêu cầu trích xuất đặc trưng thủ công từ hình ảnh và sau đó áp dụng các bài toán phân loại. Một trong những nghiên cứu tiêu biểu là "Phân loại ImageNet bằng mạng nơ-ron tích chập sâu" của Alex Krizhevsky, Ilya Sutskever và Geoffrey Hinton, đã mở ra hướng đi mới cho việc áp dụng mạng nơ-ron tích chập (CNN) trong nhận dạng hình ảnh.

Điểm hay và cải tiến của luận văn

Sử dụng mạng nơ-ron tích chập (CNN) end-to-end: Khác với các phương pháp truyền thống, luận văn này sử dụng mô hình học sâu CNN có khả năng thực hiện dự đoán end-to-end mà không cần trích xuất đặc trưng thủ công. Điều này giúp giảm thiểu công đoạn xử lý dữ liệu và tăng độ chính xác của dự đoán.

Ứng dụng thực tiễn: Một điểm nổi bật của nghiên cứu là đã xây dựng thành công một ứng dụng hỗ trợ tuyển dụng, giúp nhà tuyển dụng xác định thông tin cá nhân của ứng viên nhanh chóng và chính xác. Ứng dụng này đã được thử nghiệm tại phòng nhân sự Công ty Long Vân System Solution và cho thấy hiệu quả cao trong thực tế.

Mô hình cải tiến: Luận văn không chỉ dừng lại ở việc xây dựng mô hình CNN, mà còn tối ưu hóa mô hình để đạt hiệu quả cao nhất trong việc dự đoán độ tuổi và giới tính. Điều này bao gồm việc điều chỉnh các tham số mô hình, sử dụng các kỹ thuật tiền xử lý ảnh và tối ưu hóa quá trình huấn luyện.

Độ chính xác cao: Kết quả thử nghiệm của mô hình trên các bộ dữ liệu công khai và dữ liệu thực tế cho thấy độ chính xác cao trong việc dự đoán, với tỷ lệ chính xác đạt hơn 86% cho dự đoán giới tính và hơn 92% cho dự đoán độ tuổi. Đây là một cải tiến đáng kể so với các phương pháp truyền thống.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

1. Mô hình CNN có khả năng dự đoán độ tuổi và giới tính với độ chính xác cao, đạt được kết quả tốt trên cả bộ dữ liệu huấn luyện và kiểm tra.
2. Hệ thống hỗ trợ tuyển dụng hoạt động hiệu quả, giúp nhà tuyển dụng xác định chính xác thông tin cá nhân của ứng viên, giảm thiểu tình trạng khai man.
3. Kết quả nghiên cứu sẽ đóng góp vào việc phát triển các ứng dụng học sâu trong lĩnh vực tuyển dụng và quản lý nhân sự.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

- [1]. Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." Communications of the ACM, 2012.
- [2]. Linnan Zhu et al. "Learning a Lightweight Convolutional Neural Network for Simultaneous Age and Gender Recognition." 2017.