

KỸ THUẬT TRÍ TUỆ NHÂN TẠO

ĐỀ CƯƠNG NGHIÊN CỨU

Nhận diện chữ viết tay bằng phương pháp mạng neuron và hoc sâu

GIÁO VIÊN HƯỚNG DẪN:

1. Nguyễn Ngọc Thảo

NHÓM SV THỰC HIỆN:

1. 1512005 - Nguyễn Hoàng Anh

2. 1512031 - Huỳnh Cao Biên

Từ khóa: Neural Networks, Deep Learning, recognize handwritten digits, algorithm for intelligence, Perceptrons

Deadline: 23:59 22/03/2018



Khoa Công nghệ Thông tin
Đại học Khoa học Tự nhiên TP HCM

Tháng 3/2018

2 Contents

1.	THÔNG TIN NHÓM THỰC HIỆN.....	3
2.	NỘI DUNG THỰC HIỆN	3
2.1.	Phát biểu bài toán.....	3
2.2.	Hướng tiếp cận.....	3
2.3.	Giải pháp thực hiện.....	4
2.3.1	Giải thích Mạng nơ-ron thần kinh nhân tạo (gọi tắt là Perceptrons).....	4
2.3.2	Tế bào thần kinh sigma	4
2.3.3	Kiến trúc của mạng nơron.....	4
2.3.4	Một mạng đơn giản để phân loại chữ viết tay	4
2.3.5	Học với Gradient Descent	4
2.3.6	Hoàn thành mạng lưới để phân loại các chữ số.....	4
2.3.7	Hướng cái tiến: Học sâu – Deep Learning	4
2.4.	Dữ liệu thực nghiệm:	4
2.5.	Thực nghiệm và đánh giá:.....	4
3.	KẾ HOẠCH THỰC HIỆN.....	4
4.	TỰ ĐÁNH GIÁ NHÓM.....	5
5.	TÀI LIỆU THAM KHẢO	5

1. THÔNG TIN NHÓM THỰC HIỆN

2. NỘI DUNG THỰC HIỆN

2.1. Phát biểu bài toán

Bài toán giải quyết vấn đề gì?

Hệ thống thị giác của con người là một trong những kì quan của thế giới. Chúng ta đang tự hào rằng khả năng xử lý siêu phàm của bộ não giúp thị giác chúng ta nhìn được những gì thế giới thực đang có.

MSSV	Họ tên	Email
1512005	Nguyễn Hoàng Anh	1512005@student.hcmus.edu.vn
1512031	Huỳnh Cao Biên	1512209@student.hcmus.edu.vn

Nhìn vào chuỗi viết tay sau đây:

504192

Hầu hết mọi người dễ dàng nhận ra những chữ số như 504192. Như vậy bài toán chúng ta đặt ra là làm sao để máy tính có thể nhìn thấy và hiểu được những thứ viết tay này? Đúng là một câu hỏi lớn!

Dữ liệu đầu vào và đầu ra là gì? Output là gì?

Đĩ nhiên là những tập ảnh có chữ viết tay của con người. Output là những chữ số 0-9, chữ cái A-Z, hoặc các ký tự đặc biệt.

2.2. Hướng tiếp cận

Chúng ta hãy suy nghĩ về cách mà thị giác và bộ não của chúng ta hoạt động: Ở mỗi bán cầu não, con người có vỏ não thị giác chính, còn gọi là V1, chứa 140 triệu tế bào thần kinh, với hàng chục tỷ kết nối giữa chúng. Và thị giác của con người không chỉ liên quan đến V1, mà còn là một loạt các vỏ não thị giác - V2, V3, V4 và V5 - đang dần dần phức tạp hơn để xử lý hình ảnh. Chúng ta mang trong đầu một siêu máy tính, theo dõi bởi sự tiến hóa trong hàng trăm triệu năm, và thích nghi tuyệt vời để hiểu được thế giới thị giác. Gần như tất cả những công việc đó được thực hiện một cách vô thức.

Các giải pháp phổ biến từ trước đến nay là gì?

Ý tưởng là có một số lượng lớn chữ viết tay chữ số, được gọi là tập training. Chúng ta dùng mạng



neuron để cho máy tự động suy diễn ra các quy tắc nhận diện chữ viết tay. Hơn nữa, tang tập training lên càng cao thì độ chính xác nhận dạng sẽ càng cao. Vì vậy, trong khi tôi đã chỉ ra 100 số đào tạo ở trên, có lẽ chúng ta có thể xây dựng một công cụ nhận diện chữ viết tay tốt hơn bằng cách sử dụng hàng ngàn hoặc thậm chí hàng triệu hoặc hàng tỉ ví dụ đào tạo. Và trước giờ người ta chỉ làm cách này để nhận dạng chữ viết tay.

Vấn đề gì của bài toán đã được giải quyết?

Như thế vấn đề máy tính nhận dạng đã được giải quyết

Vấn đề gì của bài toán chưa được giải quyết?

Độ chính xác của phương pháp này không phải là tuyệt đối vì độ chính xác chúng ta đang phụ thuộc vào tập training, theo lý thuyết chúng ta có thể đạt độ chính xác tuyệt đối 100% khi và chỉ khi tập training có số lượng ở vô cùng nhưng thực tế thì không thể.

2.3. Giải pháp thực hiện

2.3.1 Giải thích Mạng nơ-ron thần kinh nhân tạo (gọi tắt là Perceptrons)

2.3.2 Tế bào thần kinh sigma

2.3.3 Kiến trúc của mạng nơron

2.3.4 Một mạng đơn giản để phân loại chữ viết tay

2.3.5 Học với Gradient Descent

2.3.6 Hoàn thành mạng lưới để phân loại các chữ số

2.3.7 Hướng cái tiến: Học sâu – Deep Learning

2.4. Dữ liệu thực nghiệm:

Tập training set gồm 60,000 mẫu. Được lấy từ 250 người. Mỗi mẫu là một ảnh 28x28 pixel.

Tập test set gồm 10,000 mẫu. Mỗi mẫu là một ảnh 28x28 pixel.

Tất cả được lấy từ MNIST data set: <http://yann.lecun.com/exdb/mnist/>

2.5. Thực nghiệm và đánh giá:

3. KẾ HOẠCH THỰC HIỆN

Tuần	Thời gian thực hiện		Công việc cụ thể
	Start	End	
6	21/3/2018	30/3/2018	Tìm hiểu đề tài và tìm kiếm tài liệu, paper có liên quan. Phát biểu bài toán, Hướng tiếp cận Chọn giải pháp, đề xuất tài liệu liên quan đến neuron network, deep learning, handwritten digits
7	30/3/2018	7/4/2018	
8	7/4/2018	14/4/2018	Giải thích Mạng nơ-ron thần kinh nhân tạo Giải thích Tế bào thần kinh sigma Giải thích Kiến trúc của mạng nơron
9	14/4/2018	21/4/2018	Trình bày Một mạng đơn giản để phân loại chữ viết tay Trình bày cách Học với Gradient Descent có liên quan đến nhận dạng Hoàn thành mạng lưới để phân loại các chữ số

10	21/4/2018	30/4/2018	<p>Tìm ra Hướng cải tiến: Học sâu – Deep Learning cho nhận dạng chữ viết</p> <p>Bắt đầu làm báo cáo chi tiết.</p> <p>Tiến hành cài đặt và chạy thực nghiệm:</p> <p>Cho hệ thống chạy trên tập training set.</p> <p>Cho hệ thống chạy thử trên tập test set và đánh giá kết quả.</p>
11	30/4/2018	7/5/2018	Hoàn tất sản phẩm và báo cáo.

4. TỰ ĐÁNH GIÁ NHÓM

5. TÀI LIỆU THAM KHẢO

[1] Neural Networks and Deep Learning :

<http://neuralnetworksanddeeplearning.com>

[2] THE MNIST DATABASE of handwritten digits :

<http://yann.lecun.com/exdb/mnist/>

[3] Gradient Descent:

<https://machinelearningcoban.com/2017/01/12/gradientdescent/>

[4] Modeling the manifolds of images of handwritten digits:

<http://ieeexplore.ieee.org/abstract/document/554192/?reload=true>