

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM
TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG
KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CUỐI KỲ MÔN NHẬP MÔN HỌC MÁY

Người hướng dẫn: **LÊ VĂN CƯỜNG**

Người thực hiện: **PHẠM NHẬT TUYÊN - 52100945**

Lớp : 21050301

Khoá : 25

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM
TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG
KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CUỐI KỲ MÔN NHẬP MÔN HỌC MÁY

Người hướng dẫn: **LÊ VĂN CƯỜNG**

Người thực hiện: **PHẠM NHẬT TUYÊN**

Lớp : **21050301**

Khoá : **25**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

LỜI CẢM ƠN

Đầu tiên, em xin gửi lời chân thành và sâu sắc đến thầy Lê Văn Cường. Trong quá trình tìm hiểu và học tập bộ môn Nhập môn học máy, em đã luôn nhận được sự giảng dạy và hướng dẫn rất tận tình, tâm huyết từ thầy. Thầy đã giúp em tích lũy thêm nhiều kiến thức hay và bổ ích để em có thể vận dụng vào bài báo cáo này.

Có lẽ bởi vì kiến thức là vô hạn mà sự tiếp nhận kiến thức lại hữu hạn. Do đó, trong quá trình hoàn thành bài tiểu luận, chắc chắn không tránh khỏi những thiếu sót. Em rất mong nhận được những ý kiến đóng góp quý báu từ các thầy cô để bài tiểu luận của em được hoàn thiện hơn.

Một lần nữa, em xin trân trọng cảm ơn!

ĐỒ ÁN ĐƯỢC HOÀN THÀNH TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG

Tôi xin cam đoan đây là sản phẩm đồ án của riêng chúng tôi và được sự hướng dẫn của Lê Văn Cường. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu **Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

TP. Hồ Chí Minh, ngày 21 tháng 11 năm 2023

Tác giả

(ký tên và ghi rõ họ tên)

Tuyên

Phạm Nhật Tuyên

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

Phần xác nhận của GV hướng dẫn

Tp. Hồ Chí Minh, ngày tháng năm
(kí và ghi họ tên)

Phần đánh giá của GV chấm bài

Tp. Hồ Chí Minh, ngày tháng năm
(kí và ghi họ tên)

MỤC LỤC

NỘI DUNG BÁO CÁO	5
1. Tìm hiểu, so sánh các phương pháp Optimizer trong huấn luyện mô hình học máy.....	5
a. Optimizer là gì?	5
b. Các phương pháp Optimizer	5
<input type="checkbox"/> Gradient Descent:	5
<input type="checkbox"/> Stochastic Gradient Descent:	5
<input type="checkbox"/> RMSprop	6
<input type="checkbox"/> Adagrad:	6
c. So sánh các thuật toán trên.....	7
2. Tìm hiểu về Continual Learning và Test Production khi xây dựng một giải pháp học máy để giải quyết một bài toán nào đó.....	8
TÀI LIỆU THAM KHẢO	9

NỘI DUNG BÁO CÁO

Bài 1: Trình bày một bài nghiên cứu, đánh giá của em về các vấn đề sau:

1. Tìm hiểu, so sánh các phương pháp Optimizer trong huấn luyện mô hình học máy

a. Optimizer là gì?

Tối ưu hóa trong học máy được coi như một phương pháp sử dụng để giảm thiểu các lỗi hoặc hàm mất mát trong quá trình huấn luyện mô hình học máy điển hình là mạng nơ-ron. Bằng cách điều chỉnh các tham số lặp đi lặp lại (trọng số - weights và độ lệch – biases) từ đó giảm thiểu các lỗi hoặc các tổn thất khác.

b. Các phương pháp Optimizer

▪ **Gradient Descent:**

Là một thuật toán tối ưu được sử dụng để đào tạo các mô hình học máy và mạng nơ-ron. Gradient Descent là phương pháp cơ bản, cập nhật tham số theo hướng ngược với gradient của hàm mất mát. Đây là phương pháp quan trọng đầu tiên và cơ bản nhất trong tối ưu hóa mô hình.

$$\theta = \theta - \alpha \cdot \nabla J(\theta)$$

Trong đó:

θ là vector chứa các tham số của mô hình cần cập nhật.

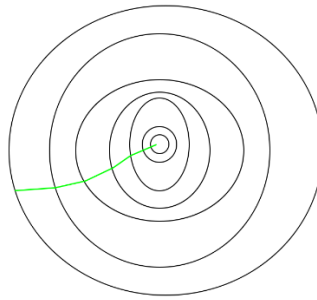
α là learning rate, quyết định tốc độ cập nhật tham số.

$\nabla J(\theta)$ là gradient của hàm mất mát J theo các tham số θ , chỉ ra hướng và độ lớn tăng nhanh nhất của hàm mất mát.

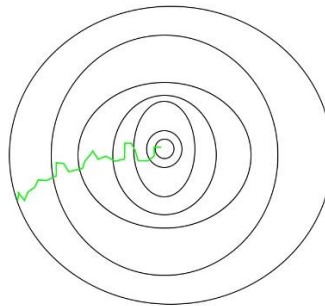
▪ **Stochastic Gradient Descent:**

Là một biến thể của thuật toán Gradient Descent. Được áp dụng cho việc tối ưu hóa mô hình học máy. Điểm đặc biệt của SGD là chỉ sử dụng một mẫu huấn luyện ngẫu nhiên hoặc một lô nhỏ dữ liệu để cập nhật mô hình, giúp giảm thiểu chi phí tính toán khi xử lý các tập dữ liệu lớn. Điều này làm tăng tính ngẫu nhiên trong quá trình tối ưu hóa, giúp tăng hiệu suất tính toán đặc biệt là khi xử lý dữ liệu lớn.

Nếu đường đi của Gradient Descent là một quỹ đạo trơn tru thì đường đi của Stochastic Gradient Descent có thể là một đường ziczac.



Hình 1: quỹ đạo của Gradient Descent



Hình 2: quỹ đạo của Stochastic Gradient Descent

▪ RMSProp

RMSProp là một thuật toán tối ưu hóa tốc độ học thích ứng trong việc huấn luyện mạng neural. Nó giảm tốc độ học dựa trên gradient bình phương của từng tham số, giúp điều chỉnh tốc độ học một cách linh hoạt hơn và giảm lượng tính toán cần thiết trong quá trình huấn luyện. RMSProp có khả năng thích ứng với từng tham số riêng biệt trong mạng, cải thiện hiệu suất so với việc sử dụng Gradient Descent thông thường.

▪ Adagrad:

Là viết tắt của Adaptive Gradient Optimizer. Giống như các thuật toán tối ưu Gradient Descent, Stochastic Gradient Descent, mini-batch SGD,

chúng đều được sử dụng để giảm hàm mất mát liên quan đến trọng số.

Công thức cập nhật trọng số như sau:

$$w_t = w_{t-1} - \eta \frac{\partial L}{\partial w(t-1)}$$

Trong đó:

$w(t)$ = giá trị của w ở lần lặp hiện tại

$w(t-1)$ = giá trị của w ở lần lặp trước

η = tốc độ học tập.

c. So sánh các thuật toán trên

Thuật toán	Ưu điểm	Hạn chế
Gradient Descent	Thường tìm được giải pháp tốt với các hàm mất mát lồi.	Tính toán gradient của hàm mất mát bằng cách sử dụng toàn bộ tập dữ liệu huấn luyện trong mỗi vòng lặp, do đó cần nhiều thời gian và tài nguyên tính toán
Stochastic Gradient Descent	Chỉ sử dụng một ví dụ ngẫu nhiên trong mỗi vòng lặp, làm cho nó nhanh hơn. Có thể thoát khỏi các điểm cực tiểu cục bộ và tìm ra giải pháp tốt hơn đối với các hàm mất mát không lồi.	Không ổn định Khó khăn trong việc đạt được sự hội tụ chính xác và có thể không tìm được giải pháp tốt với các hàm mất mát lồi.
RMSProp	giải quyết được vấn đề tốc độ học giảm dần của Adagrad	Mất cân bằng giữa các tham số: Có thể gây ra mất cân bằng giữa các tham số do việc điều chỉnh tốc độ học tùy thuộc vào gradient của từng tham số.
Adagrad	Điều chỉnh tốc độ học dựa	Adagrad đôi khi làm giảm

	trên gradient. Xử lý cấu trúc không đồng đều Hiệu quả với đặc trưng thưa	tốc độ học quá mạnh.
--	--	----------------------

2. Tìm hiểu về Continual Learning và Test Production khi xây dựng một giải pháp học máy để giải quyết một bài toán nào đó.

Continual Learning: là một quá trình học tập liên tục. Khi giải quyết các nhiệm vụ mới, bạn sử dụng kiến thức cũ để học nhanh hơn. Mục đích chính là cải thiện hiệu suất thực hiện các nhiệm vụ mới, cập nhật nền tảng kiến thức của bạn với từng trải nghiệm mới để cải tiến liên tục.

Ví dụ: Nếu bạn xây dựng một hệ thống nhận diện ảnh cho việc phát hiện các loài thú, Continual Learning cho phép mô hình cập nhật kiến thức từ dữ liệu về các loài thú mới mà nó chưa từng gặp, giúp nó nhận biết các loài thú mới một cách chính xác.

Test Production: Khi mô hình đã được huấn luyện và cập nhật thông tin, Test Production giúp tạo ra các bài kiểm tra hoặc dữ liệu kiểm tra mới từ dữ liệu thực tế để đánh giá khả năng hoạt động và hiệu suất của mô hình. Quá trình này giúp đánh giá khả năng tổng quát hóa của mô hình trước khi triển khai vào môi trường thực tế.

Ví dụ: Trong hệ thống nhận diện thú vật, Test Production tạo ra các bộ dữ liệu kiểm tra mới từ ảnh thú vật thực tế mà mô hình chưa từng gặp phải, đánh giá xem mô hình có thể nhận diện và phân loại chúng chính xác không.

TÀI LIỆU THAM KHẢO

1. <https://www.ibm.com/topics/gradient-descent#:~:text=Gradient%20descent%20is%20an%20optimization,each%20iteration%20of%20parameter%20updates.>
2. <https://www.geeksforgeeks.org/ml-stochastic-gradient-descent-sgd/>
3. <https://www.geeksforgeeks.org/gradient-descent-with-rmsprop-from-scratch/>
4. <https://blog.marketmuse.com/glossary/adaptive-moment-estimation-adam-definition/>
5. <https://www.geeksforgeeks.org/intuition-behind-adagrad-optimizer/>
6. <https://viblo.asia/p/optimizer-hieu-sau-ve-cac-thuat-toan-toi-uu-gdsgdadam-Qbq5QQ9E5D8>
7. <https://www.studocu.com/vn/document/nguyen-tat-thanh-university/cong-nghe-thong-tin/on-thi-mang-neural-va-thuat-giai/78678723>
8. https://d2l.aivivn.com/chapter_optimization/adagrad_vn.html
9. <https://www.cs.uic.edu/~liub/lifelong-machine-learning-draft.pdf> (Trang 7)
- 10.