TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN NHẬP MÔN HỌC MÁY**

**CÁC PHƯƠNG PHÁP OPTIMIZE CONTINUAL LEARNING VÀ TEST PRODUCTION**

*Người hướng dẫn*: **GV. Lê Anh Cường**

*Người thực hiện*: **Nguyễn Văn Anh - 51800009**

*Lớp*: **18050202**

*Khoá****:* 22**

**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 NHẬP MÔN HỌC MÁY**

**CÁC PHƯƠNG PHÁP OPTIMIZE CONTINUAL LEARNING VÀ TEST PRODUCTION**

*Người hướng dẫn*: **GV. Lê Anh Cường**

*Người thực hiện*: **Nguyễn Văn Anh - 51800009**

*Lớp*: **18050202**

*Khoá****:* 22**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023**

# LỜI CẢM ƠN

Để hoàn thành bài báo cáo này trước tiên nhóm chúng em xin gửi đến các quý thầy, cô giảng viên trường Đại học Tôn Đức Thắng lời cảm ơn chân thành và sâu sắc nhất. Đặc biệt, nhóm chúng em xin gửi đến thầy Lê Anh Cường – người đã tận tình hướng dẫn, giúp đỡ chúng em hoàn thành tiểu luận này lời cảm ơn sâu sắc nhất. Với điều kiện thời gian cũng như kinh nghiệm còn hạn chế của một sinh viên, bài báo cáo này không thể tránh được những thiếu sót. Chúng em rất mong nhận được sự chỉ bảo, đóng góp ý kiến của các quý thầy cô để chúng em có điều kiện bổ sung, nâng cao ý thức của mình, phục vụ tốt hơn công tác thực tế sau này.

Chúng em xin chân thành cảm ơn!.

# BÁO CÁO ĐƯỢC HOÀN THÀNH TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG

Chúng tôi xin cam đoan đây là sản phẩm đồ án của riêng chúng tôi và được sự hướng dẫn của thầy Lê Anh 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 của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**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 20 tháng 07 năm 2023*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Văn Anh*

# 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)

# TÓM TẮT

# MỤC LỤC

[LỜI CẢM ƠN i](#_Toc154264431)

[BÁO CÁO ĐƯỢC HOÀN THÀNH TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG ii](#_Toc154264432)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN iii](#_Toc154264433)

[TÓM TẮT iv](#_Toc154264434)

[MỤC LỤC v](#_Toc154264435)

[DANH MỤC HÌNH vi](#_Toc154264436)

[DANH MỤC BẢNG vii](#_Toc154264437)

[Chương 1 - TÌM HIỂU, SO SÁNH CÁC PHƯƠNG PHÁP OPTIMIZER TRONG HUẤN LUYỆN MÔ HÌNH HỌC MÁY 1](#_Toc154264438)

[1.1 Tổng quan Optimizer 1](#_Toc154264439)

[1.2 Các phương pháp Optimizer 1](#_Toc154264440)

[1.3 Kết Luận 2](#_Toc154264441)

[Chương 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. 7](#_Toc154264442)

[2.1 Continual Learning 7](#_Toc154264443)

[2.2 Test Production 7](#_Toc154264444)

[2.3 Kết luận 8](#_Toc154264445)

[Chương 3 - TÀI LIỆU THAM KHẢO 10](#_Toc154264446)

# DANH MỤC HÌNH

# DANH MỤC BẢNG

# TÌM HIỂU, SO SÁNH CÁC PHƯƠNG PHÁP OPTIMIZER TRONG HUẤN LUYỆN MÔ HÌNH HỌC MÁY

## Tổng quan Optimizer

Optimizer là một thuật toán tối ưu được sử dụng để cập nhật các tham số của mô hình học máy trong quá trình huấn luyện. Optimizer tìm cách giảm hàm mất mát của mô hình, với mục tiêu là tìm ra bộ tham số tối ưu cho mô hình.

## Các phương pháp Optimizer

Có rất nhiều phương pháp Optimizer khác nhau, mỗi phương pháp có ưu và nhược điểm riêng. Một số phương pháp Optimizer phổ biến bao gồm:

Gradient Descent (GD): Gradient Descent là phương pháp Optimizer cơ bản nhất. Phương pháp này cập nhật các tham số của mô hình theo hướng ngược lại với gradient của hàm mất mát.

Momentum: Momentum là một phương pháp Optimizer cải tiến Gradient Descent. Phương pháp này sử dụng một momentum để tăng tốc độ cập nhật các tham số của mô hình.

Adagrad: Adagrad là một phương pháp Optimizer tự động điều chỉnh tốc độ cập nhật các tham số của mô hình dựa trên độ lớn của gradient.

RMSprop: RMSprop là một phương pháp Optimizer tương tự như Adagrad, nhưng có độ ổn định cao hơn.

Adam: Adam là một phương pháp Optimizer kết hợp Momentum và RMSprop. Phương pháp này được coi là một trong những phương pháp Optimizer hiệu quả nhất hiện nay.

|  |  |  |
| --- | --- | --- |
| Phương pháp | Ưu điểm | Nhược điểm |
| Gradient Descent | Đơn giản, dễ hiểu | Tốc độ hội tụ chậm |
| Phương pháp | Ưu điểm | Có thể bị vượt biên |
| Gradient Descent | Đơn giản, dễ hiểu | Có thể bị lắc lư |
| RMSprop | Độ ổn định cao hơn Adagrad | Tốc độ hội tụ chậm hơn Adagrad |
| Adam | Hiệu quả cao | Có thể phức tạp hơn các phương pháp khác |

Lựa chọn phương pháp Optimizer phù hợp phụ thuộc vào nhiều yếu tố, bao gồm:

* Loại mô hình: Một số phương pháp Optimizer phù hợp hơn với một số loại mô hình nhất định. Ví dụ, Gradient Descent thường được sử dụng cho các mô hình đơn giản, trong khi Adam thường được sử dụng cho các mô hình phức tạp.
* Kích thước tập dữ liệu: Đối với các tập dữ liệu lớn, các phương pháp Optimizer có tốc độ hội tụ nhanh như Adam thường được ưu tiên.
* Thời gian huấn luyện: Nếu bạn có hạn chế về thời gian huấn luyện, bạn có thể cân nhắc sử dụng các phương pháp Optimizer có tốc độ hội tụ nhanh.

## Kết Luận

Optimizer là một yếu tố quan trọng trong huấn luyện mô hình học máy. Việc lựa chọn phương pháp Optimizer phù hợp có thể giúp mô hình học máy đạt được hiệu suất tốt hơn.

Optimizer là một thuật toán được sử dụng để cập nhật các tham số của mô hình học máy trong quá trình huấn luyện. Optimizer đóng một vai trò quan trọng trong việc xác định hiệu suất của mô hình.

Có nhiều phương pháp optimizer khác nhau, mỗi phương pháp có ưu và nhược điểm riêng. Dưới đây là một số phương pháp optimizer phổ biến:

* Gradient descent: Phương pháp này cập nhật các tham số theo hướng ngược của gradient của hàm mất mát. Gradient descent là phương pháp optimizer đơn giản và hiệu quả, nhưng có thể gặp khó khăn trong việc thoát khỏi các điểm tối thiểu cục bộ.
* Stochastic gradient descent: Phương pháp này cập nhật các tham số theo hướng ngược của gradient của hàm mất mát, nhưng chỉ sử dụng một phần nhỏ của tập dữ liệu trong mỗi lần cập nhật. Stochastic gradient descent có thể giúp giảm thiểu vấn đề hội tụ của gradient descent, nhưng có thể dẫn đến kết quả không nhất quán.
* Adagrad: Phương pháp này cập nhật các tham số theo hướng ngược của gradient của hàm mất mát, nhưng với tốc độ cập nhật được điều chỉnh dựa trên gradient trước đó. Adagrad có thể giúp giảm thiểu vấn đề hội tụ của gradient descent, nhưng có thể dẫn đến kết quả không ổn định.
* RMSProp: Phương pháp này tương tự như Adagrad, nhưng sử dụng bình phương của gradient trước đó để điều chỉnh tốc độ cập nhật. RMSProp có thể cải thiện độ ổn định của Adagrad.
* Adam: Phương pháp này là sự kết hợp của Adagrad và RMSProp. Adam có thể mang lại hiệu quả tốt trong nhiều trường hợp.

Một số mẹo sử dụng optimizer

Dưới đây là một số mẹo sử dụng optimizer hiệu quả:

* Khởi tạo tham số: Khởi tạo tham số của mô hình một cách hợp lý có thể giúp cải thiện hiệu quả của optimizer.
* Chọn tham số: Các phương pháp optimizer thường có một số tham số cần được điều chỉnh. Việc điều chỉnh các tham số này có thể giúp cải thiện hiệu quả của optimizer.
* Theo dõi kết quả: Theo dõi kết quả huấn luyện của mô hình có thể giúp bạn xác định xem optimizer đang hoạt động hiệu quả hay không. Nếu optimizer không hoạt động hiệu quả, bạn có thể cần thử một phương pháp optimizer khác hoặc điều chỉnh các tham số của optimizer.

Trong huấn luyện mô hình học máy, tối ưu hóa là quá trình tìm kiếm các giá trị tham số của mô hình sao cho mô hình đạt được hiệu suất tốt nhất trên tập dữ liệu huấn luyện. Các phương pháp tối ưu hóa khác nhau sẽ sử dụng các thuật toán khác nhau để tìm kiếm các giá trị tham số tối ưu.

Một số phương pháp tối ưu hóa phổ biến trong huấn luyện mô hình học máy bao gồm:

* Gradient descent (GD): GD là phương pháp tối ưu hóa phổ biến nhất trong huấn luyện mô hình học máy. Phương pháp này sử dụng đạo hàm của hàm mất mát để cập nhật các giá trị tham số của mô hình theo hướng giảm hàm mất mát.
* Stochastic gradient descent (SGD): SGD là một biến thể của GD sử dụng chỉ một mẫu dữ liệu từ tập dữ liệu huấn luyện để cập nhật các giá trị tham số của mô hình trong mỗi bước. SGD có thể hiệu quả hơn GD trong các trường hợp tập dữ liệu lớn.
* Adagrad: Adagrad là một phương pháp tối ưu hóa dựa trên GD. Phương pháp này điều chỉnh tốc độ cập nhật các giá trị tham số của mô hình dựa trên giá trị đạo hàm của hàm mất mát.
* RMSprop: RMSprop là một phương pháp tối ưu hóa dựa trên GD. Phương pháp này tương tự như Adagrad, nhưng sử dụng phương pháp trung bình động để làm mượt giá trị đạo hàm của hàm mất mát.
* Adam: Adam là một phương pháp tối ưu hóa kết hợp các ưu điểm của Adagrad và RMSprop. Adam là một phương pháp tối ưu hóa hiệu quả và được sử dụng rộng rãi trong nhiều mô hình học máy.

Lựa chọn phương pháp tối ưu hóa

Lựa chọn phương pháp tối ưu hóa phù hợp phụ thuộc vào nhiều yếu tố, bao gồm:

* Kích thước của tập dữ liệu huấn luyện
* Tính chất của hàm mất mát
* Các tính năng của mô hình

Trong các trường hợp tập dữ liệu lớn, SGD hoặc Adam là những lựa chọn tốt. Trong các trường hợp tập dữ liệu nhỏ, GD có thể là lựa chọn tốt hơn. Trong các trường hợp hàm mất mát có nhiều điểm dừng, Adagrad hoặc RMSprop có thể không hiệu quả.

Ứng dụng

Các phương pháp tối ưu hóa được sử dụng trong nhiều mô hình học máy khác nhau, bao gồm:

* Phân loại
* Hồi quy
* Tìm kiếm các mẫu
* Xử lý ngôn ngữ tự nhiên
* Học máy sâu

Ví dụ, trong các mô hình phân loại, các phương pháp tối ưu hóa được sử dụng để tìm các trọng số của mô hình sao cho mô hình có thể phân loại các mẫu dữ liệu mới một cách chính xác. Trong các mô hình hồi quy, các phương pháp tối ưu hóa được sử dụng để tìm các tham số của mô hình sao cho mô hình có thể dự đoán giá trị của một biến dựa trên các giá trị của các biến khác.

# 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

Continual Learning là một lĩnh vực của học máy tập trung vào việc học các mô hình có thể học hỏi và thích ứng với dữ liệu mới mà không bị mất hiệu suất trên dữ liệu cũ. Điều này là cần thiết trong nhiều ứng dụng thực tế, nơi dữ liệu luôn thay đổi theo thời gian.

Có nhiều phương pháp khác nhau để thực hiện Continual Learning. Một số phương pháp phổ biến bao gồm:

* Entropy regularization: Phương pháp này thêm một thuật toán entropy vào hàm mất mát của mô hình. Điều này giúp mô hình tránh tập trung quá nhiều vào các tập dữ liệu cũ.
* Data augmentation: Phương pháp này tạo ra các dữ liệu mới từ dữ liệu cũ. Điều này giúp mô hình học được các đặc trưng mới mà không cần phải học lại các đặc trưng cũ.
* Incremental learning: Phương pháp này huấn luyện mô hình mới trên dữ liệu mới. Điều này giúp mô hình học được các đặc trưng mới mà không cần phải thay đổi mô hình cũ.

## Test Production

Test Production là một quá trình trong đó một giải pháp học máy được triển khai trên môi trường sản xuất và được theo dõi hiệu suất. Điều này giúp đảm bảo rằng giải pháp học máy đang hoạt động hiệu quả và có thể đáp ứng các yêu cầu của doanh nghiệp.

Có nhiều công cụ và phương pháp khác nhau để thực hiện Test Production. Một số công cụ phổ biến bao gồm:

* Đánh giá hiệu suất: Các phương pháp đánh giá hiệu suất được sử dụng để xác định xem giải pháp học máy đang hoạt động hiệu quả hay không.
* Theo dõi hiệu suất: Các phương pháp theo dõi hiệu suất được sử dụng để giám sát hiệu suất của giải pháp học máy theo thời gian.
* Khắc phục sự cố: Các phương pháp khắc phục sự cố được sử dụng để giải quyết các vấn đề phát sinh với giải pháp học máy.

Ứng dụng của Continual Learning và Test Production

Continual Learning và Test Production có thể được áp dụng cho nhiều giải pháp học máy, chẳng hạn như:

* Dự đoán: Các giải pháp dự đoán có thể sử dụng Continual Learning để cập nhật mô hình của chúng với dữ liệu mới. Điều này giúp các giải pháp dự đoán luôn chính xác.
* Phân loại: Các giải pháp phân loại có thể sử dụng Continual Learning để học các loại dữ liệu mới. Điều này giúp các giải pháp phân loại luôn có thể phân loại các dữ liệu mới một cách chính xác.
* Hồi quy: Các giải pháp hồi quy có thể sử dụng Continual Learning để học các mối quan hệ mới giữa các biến. Điều này giúp các giải pháp hồi quy luôn có thể dự đoán các giá trị mới một cách chính xác.

Lợi ích của Continual Learning và Test Production

Continual Learning và Test Production mang lại nhiều lợi ích cho các giải pháp học máy, bao gồm:

* Hiệu suất cao hơn: Continual Learning giúp các giải pháp học máy luôn học hỏi và thích ứng với dữ liệu mới, dẫn đến hiệu suất cao hơn.
* Tính ổn định cao hơn: Test Production giúp đảm bảo rằng các giải pháp học máy đang hoạt động hiệu quả và có thể đáp ứng các yêu cầu của doanh nghiệp.
* Tính linh hoạt cao hơn: Continual Learning giúp các giải pháp học máy có thể thích ứng với các thay đổi của môi trường.

## Kết luận

Continual Learning và Test Production là hai kỹ thuật quan trọng cần được xem xét khi xây dựng một giải pháp học máy để giải quyết một bài toán nào đó. Việc áp dụng Continual Learning giúp các giải pháp học máy luôn học hỏi và thích ứng với dữ liệu mới, dẫn đến hiệu suất cao hơn. Việc thực hiện Test Production giúp đảm bảo rằng các giải pháp học máy đang hoạt động hiệu quả và có thể đáp ứng các yêu cầu của doanh nghiệp.

Có nhiều kỹ thuật Continual Learning khác nhau, mỗi kỹ thuật có ưu và nhược điểm riêng. Một số kỹ thuật Continual Learning phổ biến bao gồm:

* Forgetting prevention: Các kỹ thuật này nhằm mục đích ngăn chặn mô hình quên những gì đã học được trong quá khứ.
* Incremental learning: Các kỹ thuật này cập nhật mô hình một cách liên tục khi được cung cấp dữ liệu mới.
* Meta-learning: Các kỹ thuật này sử dụng một mô hình meta để học cách học các mô hình mới.

Test Production là một quy trình kiểm tra các mô hình học máy trong môi trường sản xuất. Quy trình này bao gồm việc thu thập dữ liệu từ môi trường sản xuất và sử dụng dữ liệu này để đánh giá hiệu suất của mô hình.

# TÀI LIỆU THAM KHẢO

**Tiếng Việt**

1. [https://www.stdio.vn/ai-ml/bai-toan-phan-lop-trong-machine-learning-classification-in-machine-learning-515OIh.](https://www.stdio.vn/ai-ml/bai-toan-phan-lop-trong-machine-learning-classification-in-machine-learning-515OIh.%20%20%20%20%20%20)  {Đã truy cập 10/12/2023 ]
2. <http://hoctructuyen123.net/tong-quan-ve-thuat-toan-phan-lop-naive-bayes-classification-nbc/>.{Đã truy cập 10/12/2023 ]
3. <https://machinelearningcoban.com/2017/08/08/nbc/>. {Đã truy cập 10/12/2023 ]
4. <https://whitehat.vn/threads/thuat-toan-phan-loai-naive-bayes-va-ung-dung.13775/> {Đã truy cập 10/12/2023 ]
5. <https://www.academia.edu/7482756/Decision_Tree> {Đã truy cập 10/12/2023 ]
6. <https://machinelearningcoban.com/2018/01/14/id3/> {Đã truy cập 10/12/2023 ]
7. <https://vietnambiz.vn/thuat-toan-k-lang-gieng-gan-nhat-k-nearest-neighbor-knn-la-gi-2020022911113334.htm> {Đã truy cập 10/12/2023 ]
8. <https://neralnetwork.wordpress.com/2018/05/11/thuat-toan-support-vector-machine-svm/> {Đã truy cập 10/12/2023 ]
9. <https://machinelearningcoban.com/2017/04/09/smv/> {Đã truy cập 10/12/2023 ]
10. [https://towardsdatascience.com/implementing-sgd- from-scratch-d425db18a72c](https://towardsdatascience.com/implementing-sgd-%20from-scratch-d425db18a72c). {Đã truy cập 10/12/2023 ]
11. <https://d2l.ai/chapter_optimization/adam.html>. {Đã truy cập 10/12/2023 ]