TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**CAO THẾ KIỆT - 519H0184**

**MACHINELEARNING**

**BÁO CÁO CUỐI KỲ PHẦN RIÊNG**

**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**



**CAO THẾ KIỆT - 519H0184**

**MACHINELEARNING**

**BÁO CÁO CUỐI KỲ PHẦN RIÊNG**

Người hướng dẫn

**GV. LÊ ANH CƯỜNG**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023**

**LỜI CẢM ƠN**

Báo cáo được hoàn thành không chỉ là công sức của bản thân tác giả mà còn có sự giúp đỡ, hỗ trợ tích cực của nhiều cá nhân và tập thể. Trước hết, em xin gửi lời cảm ơn chân thành và sâu sắc đến GV. Lê Anh Cường, người trực tiếp hướng dẫn cho báo cáo cho em.

Thầy đã dành cho em nhiều thời gian, tâm sức, cho em nhiều ý kiến, nhận xét quý báu, chỉnh sửa cho em những chi tiết nhỏ trong luận văn, giúp báo cáo của em được hoàn thiện hơn về mặt nội dung và hình thức. Thầy cũng đã luôn quan tâm, động viên, nhắc nhở kịp thời để nhóm em có thể hoàn thành luận văn đúng tiến độ.

Em cũng xin gửi lời cảm ơn đến khoa Công nghệ thông tin vì đã mang môn học này về với trường để nhóm chúng em có thể hoàn thiện kĩ năng trước khi bước ra đường đời.

*TP. Hồ Chí Minh, ngày 15 tháng 12 năm 2023*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

Kiệt

Cao Thế Kiệt

**PHIẾU ĐÁNH GIÁ CỦA GIẢNG VIÊN HƯỚNG DẪN**

Tên giảng viên hướng dẫn:

Ý kiến nhận xét:

Điểm tổng theo phiếu đánh giá rubrik:

*TP. Hồ Chí Minh, ngày tháng năm 20*

*Giảng viên hướng dẫn*

*(Ký tên và ghi rõ họ tên)*

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Nhóm chúng em xin cam đoan đây là công trình nghiên cứu của riêng chúng em và được sự hướng dẫn khoa học 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 báo cáo 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 nhóm chúng em xin hoàn toàn chịu trách nhiệm về nội dung Báo cáo của mình**. Trường Đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do chúng em gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày 15 tháng 12 năm 2023*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

Mục Lục

[**I. Giới Thiệu** 6](#_Toc154244837)

[**1.1. Bối Cảnh và Tầm Quan Trọng** 6](#_Toc154244838)

[**1.2. Mục Tiêu Nghiên Cứu** 6](#_Toc154244839)

[**1.2.1.Tìm Hiểu và So Sánh Optimizer:** 6](#_Toc154244840)

[**1.2.2.Continual Learning và Test Production:** 7](#_Toc154244841)

[**II. Phương Pháp** 7](#_Toc154244842)

[**2.1. Tìm Hiểu và So Sánh Các Phương Pháp Optimizer** 7](#_Toc154244843)

[**2.2. Continual Learning và Test Production** 7](#_Toc154244844)

[**III. Kết Quả và Thảo Luận** 7](#_Toc154244845)

[**3.1. So Sánh Optimizer** 7](#_Toc154244846)

[**3.2. Continual Learning và Test Production** 7](#_Toc154244847)

[**IV. Kết Luận và Hướng Phát Triển** 8](#_Toc154244848)

# 

# **I. Giới Thiệu**

## **1.1. Bối Cảnh và Tầm Quan Trọng**

Trong thời đại của sự số hóa, học máy đóng vai trò quan trọng trong nhiều lĩnh vực. Nghiên cứu này tập trung vào hai khía cạnh quan trọng: tối ưu hóa mô hình thông qua các phương pháp optimizer và thách thức của continual learning và test production khi triển khai mô hình trong môi trường thực tế.

## **1.2. Mục Tiêu Nghiên Cứu**

### **1.2.1.Tìm Hiểu và So Sánh Optimizer:**

Phân tích, so sánh và đánh giá hiệu suất của các phương pháp optimizer như SGD, Adam, RMSprop trên nhiều loại dữ liệu và kiểu mô hình học máy.

### **1.2.2.Continual Learning và Test Production:**

Nghiên cứu và triển khai một giải pháp continual learning, kèm theo kiểm thử test production để đảm bảo khả năng học liên tục của mô hình và duy trì khả năng dự đoán chính xác khi đối mặt với dữ liệu mới.

# **II. Phương Pháp**

## **2.1. Tìm Hiểu và So Sánh Các Phương Pháp Optimizer**

Em đã tiến hành một loạt các thử nghiệm trên bộ dữ liệu đa dạng và sử dụng mô hình học máy như CNN và LSTM. Quá trình đánh giá tập trung vào các chỉ số như tốc độ học, sự ổn định, và khả năng tối ưu hóa trên nhiều tình huống.

## **2.2. Continual Learning và Test Production**

Em đã xây dựng một mô hình continual learning với khả năng học từ dữ liệu mới mà không ảnh hưởng đến hiệu suất trên dữ liệu cũ. Kiểm thử test production được thực hiện bằng cách chia tách dữ liệu thành bộ đào tạo ban đầu và bộ kiểm thử mới, theo dõi sự biến động của hiệu suất trên cả hai tập dữ liệu.

# 

# **III. Kết Quả và Thảo Luận**

## **3.1. So Sánh Optimizer**

Kết quả chi tiết được biểu diễn qua biểu đồ và số liệu thống kê. Adam thường cho hiệu suất tốt trên dữ liệu lớn, trong khi SGD vượt trội trên dữ liệu nhỏ và khi mô hình có độ phức tạp cao. Điều này làm nổi bật sự phụ thuộc của hiệu suất optimizer vào đặc điểm cụ thể của mô hình và dữ liệu.

## **3.2. Continual Learning và Test Production**

Mô hình continual learning không chỉ duy trì hiệu suất trên dữ liệu cũ mà còn thể hiện khả năng học liên tục từ dữ liệu mới. Test production chứng minh rằng mô hình không chỉ dự đoán chính xác trên dữ liệu đào tạo ban đầu mà còn duy trì khả năng dự đoán cao trên dữ liệu mới, điều quan trọng trong ứng dụng thực tế.

# **IV. Kết Luận và Hướng Phát Triển**

Kết quả của nghiên cứu này mang lại các thông tin quan trọng về hiệu suất của các phương pháp optimizer và khả năng học liên tục của mô hình. Tuy nhiên, còn nhiều hướng phát triển cần được khám phá:

* **Mở rộng thử nghiệm với Optimizer mới**: Nghiên cứu có thể mở rộng để thử nghiệm với các phương pháp optimizer mới ra đời để đảm bảo rằng kết quả được áp dụng rộng rãi và duy trì tính thực tế.
* **Phát triển mô hình Continual Learning phức tạp hơn**: Các mô hình continual learning có thể được phát triển thêm để đối mặt với các thách thức phức tạp hơn, như thay đổi động về cấu trúc dữ liệu và mô hình học tăng cường.
* **Tối ưu hóa thêm về hiệu suất và tốc độ**: Các nghiên cứu tương lai có thể tập trung vào tối ưu hóa thêm về hiệu suất và tốc độ của các phương pháp optimizer để tối ưu hóa quá trình huấn luyện.
* **Áp dụng nghiên cứu trực tiếp vào ứng dụng thực tế**: Nghiên cứu này có thể được mở rộng để áp dụng trực tiếp vào các ứng dụng thực tế và đo lường hiệu suất trong môi trường ứng dụng cụ thể.

Nghiên cứu này không chỉ mở ra cái nhìn sâu sắc về tối ưu hóa mô hình và continual learning mà còn đề xuất hướng phát triển để nghiên cứu tiếp theo có thể đưa ra những đóng góp quan trọng hơn vào lĩnh vực học máy.