TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**FINAL PROJECT MACHINE LEARNING**

**Final Project**

*Người hướng dẫn*: **GV LÊ ANH CƯỜNG**

*Người thực hiện*: **GIANG HOẰNG ĐẠT – 521H0498**

Lớp **: 21H50301**

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**



**FINAL PROJECT MACHINE LEARNING**

**Final Project**

*Người hướng dẫn*: **GV LÊ ANH CƯỜNG**

*Người thực hiện*: **GIANG HOẰNG ĐẠT – 521H0498**

Lớp **: 21H50301**

Khoá  **: 25**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023**

LỜI CẢM ƠN

I extend my sincere appreciation to my machine learning instructor for the invaluable guidance and support provided throughout the completion of my thesis. Their expertise and dedication to teaching have been instrumental in enhancing my understanding of this intricate subject. I am grateful for the time and encouragement extended, which significantly contributed to the successful culmination of my academic work. Their passion for the field and commitment to fostering learning has been a constant source of inspiration, and I am truly fortunate to have benefited from their teaching. Thank you for your unwavering support and dedication to my academic journey in the realm of machine learning.

# ĐỒ ÁN ĐƯỢC HOÀN THÀNH

# TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG

Tôi xin cam đoan đây là sản phẩm đồ án của riêng tôi / chúng tôi và được sự hướng dẫn của GV 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 tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Giang Hoằng Đạt*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

**Phần xác nhận của GV hướng dẫn**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày 24 tháng 12 năm 2023

(kí và ghi họ tên)

**Phần đánh giá của GV chấm bài**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày 24 tháng 12 năm 2023

(kí và ghi họ tên)

CONTENTS

**CONTENTS**

[LỜI CẢM ƠN i](#_Toc154064230)

[ĐỒ ÁN ĐƯỢC HOÀN THÀNH ii](#_Toc154064231)

[TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG ii](#_Toc154064232)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN iii](#_Toc154064233)

[CHAPTER I – OPTIMIZER IN MACHINE LEARNING 2](#_Toc154064234)

[1. Overview of Machine Learning Optimization Techniques: 2](#_Toc154064235)

[2. Optimization Methods: 2](#_Toc154064236)

[2.1 Gradient Descent (GD): 2](#_Toc154064237)

[2.2 Stochastic Gradient Descent (SGD): 3](#_Toc154064238)

[2.3 Mini-Batch Gradient Descent: 3](#_Toc154064239)

[2.4 Momentum: 3](#_Toc154064240)

[2.5 Adagrad: 4](#_Toc154064241)

[2.6 RMSprop: 4](#_Toc154064242)

[2.7 Adam: 4](#_Toc154064243)

[2.8 Nadam: 5](#_Toc154064244)

[2.9 Adadelta: 5](#_Toc154064245)

[CHAPTER II – CONTINUAL LEARNING AND TEST PRODUCTION 6](#_Toc154064246)

[1. Continual Learning: 6](#_Toc154064247)

[1.1 Significance in Machine Learning Solutions: 6](#_Toc154064248)

[1.2 Challenges and Solutions: 6](#_Toc154064249)

[2. Test Production: 7](#_Toc154064250)

[2.1 Significance in Machine Learning Solutions: 7](#_Toc154064251)

[2.2 Significance in Machine Learning Solutions: 8](#_Toc154064252)

CHAPTER I – OPTIMIZER IN MACHINE LEARNING

1. Overview of Machine Learning Optimization Techniques:

The primary aim of machine learning is to develop a model that excels in making accurate predictions within a specified set of scenarios. Achieving this objective involves the crucial step of machine learning optimization.

Machine learning optimization is the focused task of adjusting hyperparameters to minimize the cost function, which gauges the difference between actual and predicted values.

2. Optimization Methods:

Optimization methods are a set of techniques and algorithms designed to find the best possible solution to a problem from a set of feasible solutions. These methods are widely used in various fields, including mathematics, engineering, computer science, economics, and many others. The primary goal of optimization is to maximize or minimize a certain objective function while satisfying a set of constraints.

2.1 Gradient Descent (GD):

Gradient Descent is an iterative optimization algorithm used for finding the minimum of a function. It updates the parameters of a model in the direction opposite to the gradient of the loss function with respect to the parameters.

* Pros:

Simple and easy to implement.

Effective for large-scale problems with extensive data.

* Cons:

Inefficient for problems with numerous parameters and non-normalized data.

2.2 Stochastic Gradient Descent (SGD):

Stochastic Gradient Descent is an optimization algorithm that randomly selects a subset of data for each iteration to compute the gradient and update model parameters.

* Pros:

Efficient for large datasets, reducing complexity compared to GD.

* Cons:

More unstable than GD due to weight updates after each data sample.

2.3 Mini-Batch Gradient Descent:

Mini-Batch Gradient Descent is a compromise between GD and SGD. It uses a small, randomly selected subset of data for each iteration.

* Pros:

Combines advantages of GD and SGD, reducing complexity.

* Cons:

Requires appropriate batch size selection depending on the problem.

2.4 Momentum:

Momentum is an optimization technique that adds a fraction of the past gradient to the current update, which helps accelerate convergence and navigate through local optima.

* Pros:

Overcomes local optimization challenges, accelerates convergence.

* Cons:

Requires proper tuning of hyperparameters to avoid overshooting.

2.5 Adagrad:

Adagrad is an adaptive learning rate optimization algorithm that adjusts the learning rates of model parameters based on the historical gradients.

* Pros:

Automatically adjusts learning rates based on gradient history.

* Cons:

Learning rate decreases rapidly, may lead to premature convergence.

2.6 RMSprop:

RMSprop is an adaptive learning rate optimization algorithm that addresses Adagrad's rapid learning rate decay by using a moving average of squared gradients.

* Pros:

Addresses Adagrad's issues by reducing the learning rate decay.

* Cons:

May still encounter challenges related to learning rate.

2.7 Adam:

Adam combines the ideas of Momentum and RMSprop. It maintains moving averages of both gradients and squared gradients and adapts the learning rate for each parameter.

* Pros:

Combines benefits of Momentum and RMSprop, often works well across diverse problems.

* Cons:

Requires tuning of multiple hyperparameters, potentially complex for small problems.

2.8 Nadam:

Nadam is an extension of Adam that incorporates Nesterov accelerated gradients. It combines Adam's adaptive learning rate with Nesterov Momentum.

* Pros:

Integrates features of both Adam and Nesterov Momentum, often exhibits high performance.

* Cons:

Still demands attention to multiple hyperparameters.

2.9 Adadelta:

Adadelta is an extension of Adagrad that aims to solve its drawback of a monotonically decreasing learning rate by using a moving average of squared parameter updates.

* Pros:

Mitigates Adagrad's learning rate problem, does not require an initial learning rate setting.

* Cons:

May necessitate the selection of specific hyperparameters.

CHAPTER II – CONTINUAL LEARNING AND TEST PRODUCTION

1. Continual Learning:

Continual learning is an approach in machine learning where a model learns sequentially from a stream of data, adapting to new information without forgetting previously acquired knowledge. It addresses the challenge of learning from non-stationary and evolving distributions over time.

* 1. Significance in Machine Learning Solutions:
* **Adapting to Dynamic Environments:**

In real-world applications, data distributions may change over time. Continual learning allows models to adapt to these changes and incorporate new knowledge without retraining on the entire dataset.

* **Resource Efficiency:**

Instead of training models from scratch whenever new data arrives, continual learning facilitates incremental updates, saving computational resources.

### 1.2 ****Challenges and Solutions:****

* **Catastrophic Forgetting:**

**Models may forget previously learned tasks when trained on new ones. Techniques like Elastic Weight Consolidation (EWC) and replay strategies address this issue by preserving important parameters.**

* **Task Boundaries:**

**Recognizing task boundaries is crucial. Methods like Task-Agnostic Online Learning (Tao) help models identify when they are transitioning to a new task.**

* **Memory Management:**

**Efficient handling of memory is essential. Techniques like Experience Replay involve periodically revisiting and training on past experiences.**

2. Test Production:

In machine learning, test production involves creating and implementing test procedures to evaluate the performance, accuracy, and reliability of a machine learning solution.

2.1 Significance in Machine Learning Solutions:

* **Quality Assurance:**

Testing ensures that the machine learning model functions as expected, producing reliable and accurate predictions.

* **Identifying Weaknesses:**

Through testing, weaknesses, biases, and limitations of the model can be identified and addressed.

* **Verification of Requirements:**

Tests verify that the model meets the specified requirements and objectives.

2.2 Significance in Machine Learning Solutions:

* Test Coverage:

Ensuring comprehensive coverage of various aspects of the model can be challenging.

* **Resource and Time Constraints:**

Limited resources and tight deadlines may impact the thoroughness of testing.

* **Changing Requirements:**

If requirements change frequently, test scenarios and cases may need continuous updates.