VIETNAM GENERAL CONFEDERATION OF LABOR
TON DUC THANG UNIVERSITY
FACULTY OF INFORMATION TECHNOLOGY

TRUONG ANH HI - 520H0356

REPORT
MACHINE LEARNING

HO CHI MINH CITY,2023

VIETNAM GENERAL CONFEDERATION OF LABOR
TON DUC THANG UNIVERSITY
FACULTY OF INFORMATION TECHNOLOGY

TRUONG ANH HI- 520H0356

REPORT
MACHINE LEARNING

Instructor:
PGS.TS.Lê Anh Cường

HO CHI MINH CITY, 2023

ACKNOWLEDGEMENT1

I am grateful for the opportunity to explore the fascinating world of machine learning and share our findings with you. We hope this report will serve as a valuable resource for those interested in learning more about this rapidly evolving field.

Thank you once again for your support.

Sincerely,

Ho Chi Minh City, 20 December 2023

Author

(Signature and full name)

Truong Anh Hi

ACKNOWLEDGEMENT2

PROJECT COMPLETED

AT TON DUC THANG UNIVERSITY

We hereby confirm that this research project is our own work and was conducted under the scientific guidance of PGS.TS.Lê Anh Cường. The research contents and results in this study are truthful and have not been previously published in any form. The data in the tables used for analysis, which are clearly cited in the reference section.

Furthermore, this project also incorporates some observations, evaluations, and data from other authors and organizations, all of which are cited and referenced accordingly.

If any form of misconduct is detected, we take full responsibility for the content of our project. Ton Duc Thang University is not responsible for any copyright infringement of violations caused by us during the course of this project (if any)

Ho Chi Minh City, 20 December 2023

Author

(Signature and full name)

Truong Anh Hi

TABLE OF CONTENTS

QUESTION 1. Optimizers for Machine Learning

1.1 Optimizers for Machine Learning

In machine learning, an optimizer is an algorithm used to update the parameters of a model during training. The optimizer finds the optimal set of parameters that allows the model to achieve the best performance on the training data.

There are many different optimizers, each with its own advantages and disadvantages. Some common optimizers include:

- **Gradient descent**: This method updates the parameters in the direction of the negative gradient of the objective function. The gradient is the derivative of the objective function at the current point of the model.
- **Stochastic gradient descent**: This method updates the parameters based on the gradient of a single random data point. This method is often faster than gradient descent, but it may not be as effective in some cases.
- **Adagrad**: This method adjusts the size of the update step based on the magnitude of the gradient. This method can help the model avoid getting stuck in local minima.
- **RMSProp:** This method is similar to Adagrad, but it uses the moving average of the gradient to adjust the size of the update step. This method can help the model be more stable during training.
- **Adam**: This method combines the advantages of Adagrad and RMSProp. This method is one of the most popular optimizers today.

1.2 Comparison of Optimizers

The following table compares the common optimizers:

Method	Advantage	Disadvantage
Gradient descent	Simple, easy to understand	Can get stuck in local minima
Stochastic gradient descent	Faster than gradient descent	May not be as effective in some cases

Adagrad	Can help the model avoid getting stuck in local minima	Can be too sensitive to data points with large gradients
RMSProp	Can help the model be more stable during training	Can be too sensitive to data points with large gradients
Adam	Combines the advantages of Adagrad and RMSProp	Can be more complex than other optimizers

1.3 Choosing an Optimizer

The choice of optimizer depends on many factors, including:

- **Type of machine learning model**: Some optimizers are better suited for certain types of machine learning models. For example, the Adam optimizer is often used for deep learning models.
- **Size of the training dataset**: Stochastic gradient descent is often a good choice for large datasets.
- **Training speed**: If we need to train the model quickly, we can use stochastic gradient descent or Adam.

QUESTION 2. Continual Learning and Test Production

2.1 Continual Learning

Continual learning is a research area in machine learning that focuses on training machine learning models that can learn and adapt to new data without forgetting what they have already learned.

In real-world applications, data often changes over time. For example, in the field of image recognition, product labels may change. If a machine learning model

is trained on a dataset of old product labels, the model may not be able to classify new products with new labels accurately.

Continual learning can be used to solve this problem. Continual learning methods help machine learning models learn and adapt to new data without forgetting what they have already learned.

For intance:

Suppose we have a machine learning model trained on a dataset of 10,000 product images. Product labels may change over time. For example, a new label may be added.

If we do not use continual learning, our machine learning model may not be able to accurately classify new products with new labels.

With continual learning, our machine learning model can continue to learn and adapt to new data. This can help our model classify new products with new labels accurately.

2.2 Test Production

Test production is the process of deploying a machine learning model into production while still continuing to train the model on new data.

Test production can be used to combine continual learning with production systems. When a machine learning model is deployed into production, the model can continue to learn and adapt to new data. This can help improve the performance of the model over time.

For instance:

Suppose we have a machine learning model trained on a dataset of 10,000 data points. This model is deployed into production and used to classify products.

With test production, the machine learning model can continue to learn and adapt to new data without being disconnected from the production system. This can help improve the performance of the model over time.