

Daily Report

Report Day: 14/2/2025

Name: Lê Quang Hưng

Team: OJT Team 4

Team Leader: Mai Tuấn Anh

Team Members: Mai Tuấn Anh, Bùi Xuân Hoàng, Lê Quang Hưng



Task Title: Exporting a PyTorch Model to ONNX

1. Purpose

Convert a trained Pytorch model to the ONNX format and compare inference performance

2. Action

- Read the documentation "**Export a PyTorch model to ONNX**".
- Trained a simple model on the MNIST dataset.
- Measured inference time using PyTorch.
- Exported the trained model to ONNX.
- Ran inference using ONNX Runtime and compared execution time.

3. Result

- Successfully exported the model to ONNX format.
- PyTorch inference time: 0.031522 seconds (CPU) and 0.021080 seconds (GPU)
- ONNX inference time: 0.008102 seconds (CPU) and 0.006233 seconds (GPU)
- ONNX execution was faster than PyTorch

