

Assignment 3: End-to-End Hugging Face Model Training & Docker Deployment

Objective

In this assignment, you will build a complete machine learning workflow starting from a notebook, converting it into production-ready scripts, training and evaluating a model using Hugging Face tools, containerizing the workflow using Docker, and publishing artifacts to GitHub and your Hugging Face profile.

Prerequisites

- Python
- Basic Machine Learning concepts
- Docker fundamentals
- Git and GitHub usage
- Basic understanding of Hugging Face ecosystem

Assignment Tasks

Task 1: Download Shared Notebook

Link:

<https://colab.research.google.com/drive/1-DhcPi4j3VBVFt9K39e895aIkXyNLJwS?usp=sharing>

Download the [ML_DL_Ops_Ass_3-Fine-Tuning-Classification.ipynb](#) file provided by the instructor.

1. Place it inside your project working directory.

Task 2: Create Your Environment Using Docker

2. Write a Dockerfile using an appropriate base image.
3. Install required dependencies.
4. Set up working directory.
5. Build and run your container.
6. Verify Python and required libraries are working inside the container.

Task 3: Convert Notebook to Python Scripts

7. Convert the downloaded notebook to .py script(s).
8. Clean unnecessary notebook artifacts.
9. Organize scripts into modules (train, eval, data, utils).

Task 4: Load Models from Hugging Face

10. Select a suitable pre-trained model (can be the same as per ipynb file).
11. Load tokenizer and model.

12. Document why you selected that model.

Task 5: Train Model Using Trainer API

13. Prepare dataset.
14. Configure training arguments.
15. Train using Trainer API.
16. Log training metrics.

Task 6: Evaluate Model

17. Run evaluation on validation/test data.
18. Record Accuracy / F1 / Loss (as applicable).
19. Save evaluation results.

Task 7: Save Model to Your Hugging Face Profile

20. Create or use your Hugging Face account.
21. Push model, tokenizer, and training config.
22. Ensure model is publicly accessible.

Task 8: Re-evaluate Model from Your Hugging Face Repo

23. Load model from your uploaded repository.
24. Run evaluation again.
25. Compare metrics with local model evaluation.

Task 9 : Create Final Docker Image (Evaluation Only)

26. Create production Docker image that pulls model from your Hugging Face profile.
27. Run evaluation automatically inside container on startup.

Task 10: Push Everything to GitHub

28. Push source code, Dockerfile, requirements file, and README.
29. Include evaluation results and model link.

Submission Requirements

- GitHub Repository Link
- Hugging Face Model Link
- Docker Image Build Instructions
- Short Report (explaining model selection, training summary, evaluation comparison, challenges)