

AI Model Fine-Tuning Project

Applied AI workflow for adapting a compact language model to phishing triage and short analyst summaries.

Aim

The point of this project is not to say "I use AI". It is to show that I can structure a real fine-tuning workflow:

- define a use case
- format data
- fine-tune a base model
- evaluate output quality
- review failure cases

What is in this project

- `data/sample_train.jsonl` : small sample training data
- `scripts/train_lora.py` : LoRA fine-tuning workflow
- `scripts/evaluate.py` : simple evaluation script

Use case

Input:

- email or message content
- metadata if available

Output:

- phishing or benign label
- short analyst-facing summary

Why it belongs in the portfolio

It shows applied AI work in a form that is concrete enough to discuss in an interview: dataset structure, prompt formatting, parameter-efficient fine-tuning, and evaluation.