National College of Ireland

Configuration Manual

MSc Research Practicum
MSc in Artificial Intelligence

Bharath Reddy Choudary
Student ID: X23353414

School of Computing
National College of Ireland

Supervisor: Mohit Garg

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Bharath Reddy Choudary

Student ID: X23353414

Programme: MSc in Artificial Intelligence **Year:** 2024-2025

Module: MSc in Research Practicum

Lecturer: Mohit Garg

Submission Due Date: 15/09/2025

Project Title: Enhancing Phishing Detection Using O3-Mini Chain-of-Thought Reasoning with GPT-4o

Word Count: **502** **Page Count: 6**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Bharath Reddy Choudary

Date: 14/09/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

1. Project Overview

This code implements an advanced phishing detection system that uses o3-mini to generate chain-of-thought reasoning for email classification, then fine-tunes GPT-4o models with this augmented data. The system processes the Enron email dataset, generates comprehensive reasoning chains, performs fine-tuning, and conducts comparative evaluations between baseline, few-shot, fine-tuned, and o3-mini models.

2. Google Colab Setup Guide

Step-by-step instructions:

1. Initial Setup

- Open Google Colab (colab.research.google.com)
- Upload o3mini_phishing_detection.ipynb or create new notebook
- Connect to T4 GPU: Runtime → Change runtime type → Hardware accelerator → T4 GPU
- Mount Google Drive when prompted for persistent storage

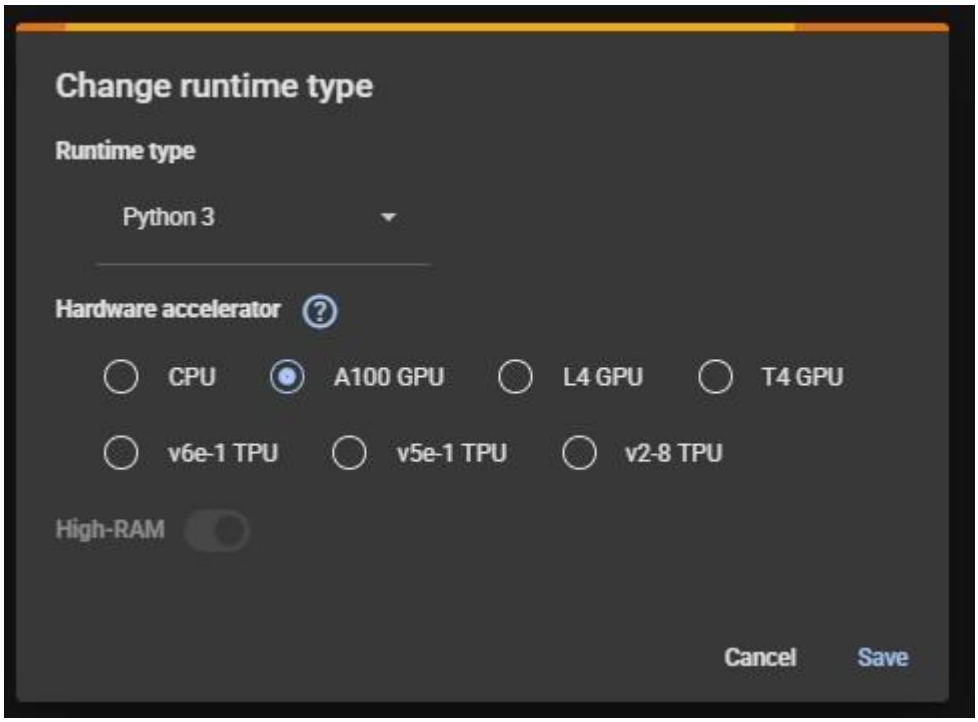

Figure 1 Run time selection

2. Dependencies & Environment

Install all required packages:

!pip install openai tiktoken pandas numpy matplotlib seaborn plotly wordcloud nltk textstat textblob scikit-learn -q

API Keys Required:

- OpenAI API key with o3-mini and GPT-4o access
- Obtain from: https://platform.openai.com/api-keys **3. Critical Configuration Points**

API Key Configuration:

API_KEY = "sk-proj-YOUR_ACTUAL_API_KEY_HERE" # <- *CHANGE THIS*

```
104         self._create_output_directories()
105         self._setup_logging()
106
107         # API configuration
108         if api_key:
109             self.client = OpenAI(api_key=api_key)
110         else:
111             # Default API key from original code
112             api_key = "sk-proj-TBa0wOtRrItPwHGSXUp_DvwfYxdpzBSxbVd_bVCmluuT2_d8l0PON5p
113             self.client = OpenAI(api_key=api_key)
114
```

Figure 2 API configuration

Dataset Path Configuration:

DATASET_PATH = "Enron.csv" # <- *CHANGE THIS if dataset is in different location*

```
46
47  class O3MiniOptimizedPipeline:
48      """O3-Mini optimized pipeline with enhanced performance and quality selection"""
49
50      def __init__(self,
51                   dataset_path="Enron.csv",
52                   samples_per_class=5000,
53                   output_dir="output",
54                   gdrive_base_path="/content/drive/MyDrive/phishing_detection_final",#
55                   checkpoint_interval=50,
56                   max_concurrent=8,
57                   api_key=None,
58                   o3_speed_mode="medium"):   # low,medium,high
59          """Initialize the O3-Mini speed-optimized pipeline"""
```

Figure 3 dataset path

Google Drive Base Path:

gdrive_base_path="/content/drive/MyDrive/phishing_detection_final" # <- *CHANGE THIS*

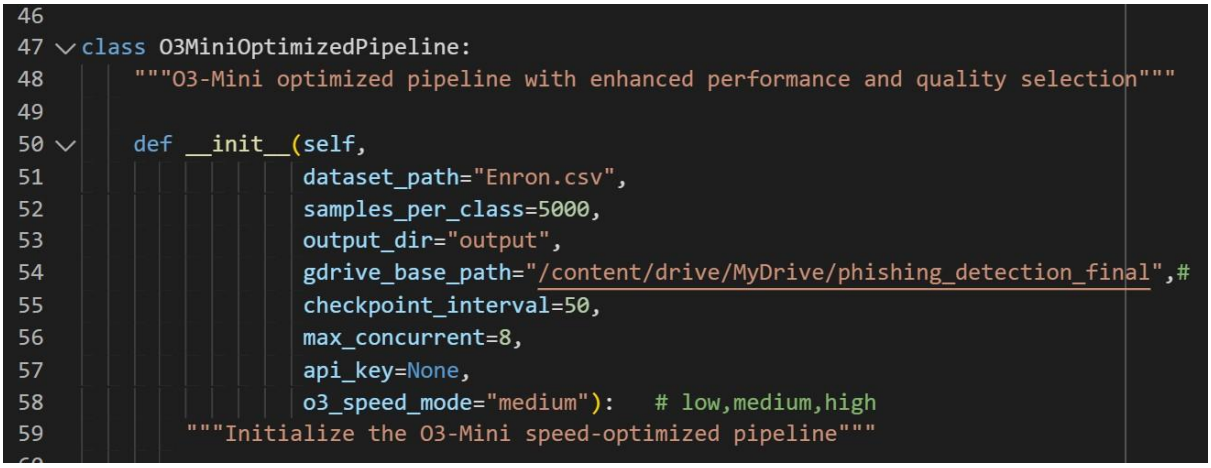

Figure 4 google drive path

Model Names (After Fine-tuning):

FINE_TUNED_MODEL = "ft:gpt-4o-mini-2024-07-18:personal:phishingdetection:YOUR_ID" # <- *CHANGE THIS*

Processing Parameters: samples_per_class=5000 # <- *CHANGE THIS*

(reduce for testing, e.g., 100)

Line 78: Adjust concurrent processing max_concurrent=8 # <-

CHANGE THIS (reduce to 4-6 if rate limited) **O3-Mini**

Configuration:

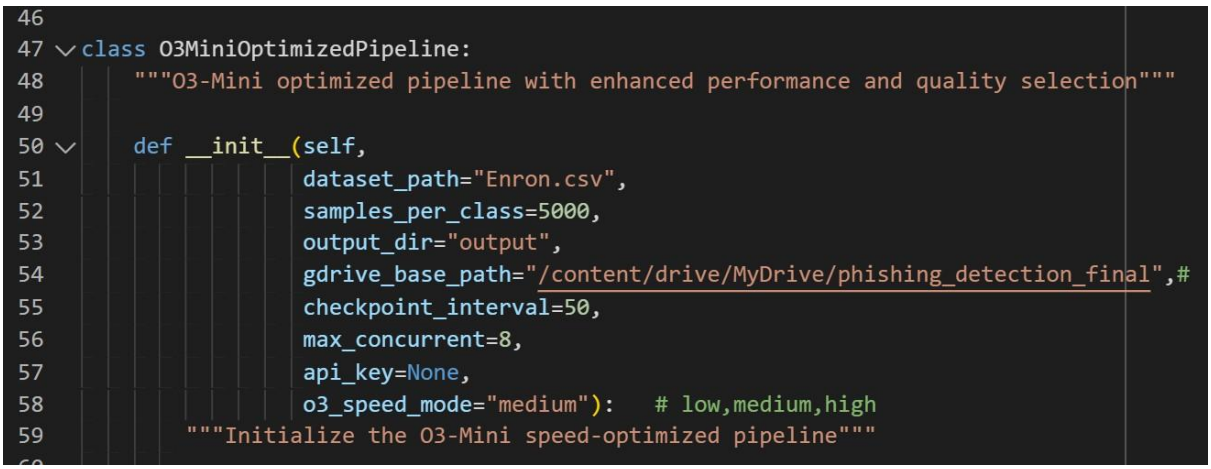

Figure 5 reasoning mode update

Line 77: Set reasoning effort level o3_speed_mode="medium" # <- *CHANGE THIS to "low" for speed or "high" for accuracy*

3. Execution Workflow

Phase 1: O3-Mini Reasoning Generation

Test with small sample first

run_o3_mini_speed_test(samples_per_class=10, speed_mode="medium")

 # Full dataset processing run_o3_mini_optimized_pipeline(samples_per_class=5000, speed_mode="medium")

Phase 2: Data Analysis

Update dataset path to your generated file

DATASET_PATH = "/content/drive/MyDrive/phishing_detection_final/output/final_datasets/o3_mini_optimized_dataset_TIMESTAMP.csv" *# <- CHANGE THIS*

Figure 6 reasoning enhanced dataset path

Phase 3: GPT-4o Fine-tuning

DATASET_PATH = "/content/drive/MyDrive/phishing_detection_final/output/final_datasets/o3_mini_optimized_dataset_TIMESTAMP.csv" *# <- CHANGE THIS*

Figure 7 reasoning enhanced dataset path

Phase 4: Model Evaluation

DATASET_PATH = "/content/drive/MyDrive/phishing_detection_final/output/final_datasets/o3_mini_optimized_dataset_TIMESTAMP.csv" *# <- CHANGE THIS*

4. File Structure Requirements

Create this directory structure in Google Drive:

/MyDrive/

└── phishing_detection_final/

├── output/

│ ├── final_datasets/

│ ├── intermediate/

```
│   ├── logs/
│   ├── checkpoints/
│   └── response_logs/
└── Enron.csv  # Upload dataset here
```

5. Data Files

Required Input:

- Enron.csv - Main dataset with columns: subject, body, label (0=legitimate, 1=phishing)

Generated Outputs:

- o3_mini_optimized_dataset_*.csv - Enhanced dataset with reasoning chains
- gpt4o_mini_training_*.jsonl - Fine-tuning training data
- model_evaluation_results_*.json - Comparison results **6. Troubleshooting**

Common Issues:

. **Rate Limiting:** *# Reduce concurrent requests*

max_concurrent=4 *# Lower from default 8*

. **Memory Issues:** *# Process smaller batches*

samples_per_class=1000 *# Reduce from 5000*

. **API Errors:**

 # Verify API key has required permissions

 # Check quota at: https://platform.openai.com/usage

7. Quick Start Commands

Test Setup (100 emails):

 # Quick test to verify everything works run_o3_mini_speed_test(samples_per_class=50, speed_mode="low")

Production Run:

Full pipeline execution run_complete_pipeline()

For fine-tuning run_colab_1000_comparison()

For model comparison

8. Expected Runtime

- **O3-Mini Processing:** ~3-5 hours for 10,000 emails
- **Fine-tuning:** 30-60 minutes
- **Model Comparison:** ~2 hours for 1000 emails