University of East London

Assignment Brief CN7050 (Intelligent Systems):

Turn it in Submission:

Class ID : 8018386

Enrolment Key : CN7050

Please read your assignment brief carefully in order to complete and submit the required components at the specified times and in the indicated formats. In case of any questions, please clarify with your module lecturer/PgDipAI Programme Lead.

Assessment Overview

This module is based on 100% continuous assessment. The assessment consists of TWO components: 20% of in-class exercise or quizzes and 80% of coursework.

A student must attend to all the components of the assignments.

Each student may be allocated to a group (of 3-4 students per group). The list of groups will be released by Week 6. The lecturer retains the right and discretion to deduct marks from individual students where group members provide evidence that the student has not contributed adequately to the assignment.

In-class Exercise or Quizzes (20%)

A total of two in-class quizzes will be conducted during the course based on the core lectures and readings supplied during the class and in the Moodle. Each student must attempt all quizzes to gain maximum marks in this component. **The quizzes will be available on Moodle in Week 8**.

Pass mark is 50% (~10).

Deliverable	Marks	Due Date
2 In-class quizzes	20% (10% + 10%)	Week 12

Coursework

Assignment Title: Analysing Financial Sentiment through Fine-tuned Transformers and Reasoning-based Large Language Models.

Weighting: 80%

Pass mark is 50% (~40)

Dataset: Sentiment Analysis for Financial News (Kaggle: ankurzing).

Submission: Moodle / Turnitin (Jupyter or Colab notebook + written report + presentation slides).

Overall mark for coursework comes from two main deliverables:

University of East London

Submission Pack	Marks
A PDF Report (around 3,000 words) along with its source python notebooks/scripts as supporting materials. Follow the example naming conventions below: TEAM1_CN7050_Report.pdf TEAM1_ CN7050_ModelA.ipynb/.py TEAM1_ CN7050_ModelB.ipynb/.py TEAM1_ CN7050_ModelC.ipynb/.py	60%
Presentation (.pptx) and (Online video presentation will be recorded) Follow the example naming conventions below: TEAM1_CN7050_dissimination.ppt	20%

Background

In the financial sector, understanding market sentiment from news headlines is a key part of intelligent decision-making. Modern NLP models such as BERT and FinBERT have transformed the way textual data can be converted into actionable insights.

While BERT-base is trained on general English text, FinBERT has been pre-trained specifically on financial reports and news articles, capturing the domain's linguistic nuances.

This assignment allows you to explore how domain-specific pretraining affects downstream performance by fine-tuning both models on a labelled financial sentiment dataset and comparing their results across multiple metrics.

You will also experiment with GPT-OSS (20B) to generate reasoning-based sentiment predictions, allowing you to compare explicit reasoning outputs with learned model predictions.

A sample code notebook for GPT-OSS will be provided in the Assignment section on Moodle.

By completing this assignment, students will be able to:

- Apply Transformer-based NLP architectures (BERT and FinBERT) for sentiment classification in a financial context.
- Compare and evaluate domain-specific versus general pretraining in fine-tuned models.
- Design and implement a predictive NLP system aligned with intelligent decision-making in FinTech.
- Critically interpret results, explain misclassifications, and discuss ethical implications in automated financial sentiment analysis.
- Integrate insights from GPT-OSS reasoning as an advanced extension for evaluation, explanation, and reflection on large language model reasoning.

Tasks and Deliverables

Part 1: Model Fine-Tuning and Evaluation (40%)

Data – Data Preparation

- Load the dataset containing financial news headlines and sentiment labels.
- Clean and normalise the data (remove duplicates, fix label casing, handle missing values).
- Split the data into training, validation, and testing subsets.
- Clearly present dataset distribution and preprocessing steps using appropriate visualisations in the report.

Model A – BERT-base

- Fine-tune **BERT-base-uncased** ([link](#)) on the dataset.
- Evaluate using macro-F1, per-class F1, accuracy, and confusion matrix.
- Visualise training and validation loss curves and discuss model convergence.

Model B – FinBERT

- Fine-tune **ProsusAI/finbert** ([link](#)) on the same dataset.
- Repeat the same training, evaluation, and visualisation pipeline as BERT.
- Compare results with Model A and interpret the differences in performance.

Expected Insight:

FinBERT is expected to achieve higher performance due to its domain-specific pretraining. Students should discuss *why* this happens, linking observations to linguistic familiarity, contextual embeddings, and feature representation quality.

Model C – GPT-OSS Reasoning Extension

Use the GPT-OSS (20B) model to demonstrate reasoning-based sentiment analysis on a small subset of 80 headlines from the test set.

- Design and apply a prompt that guides GPT-OSS to reason step-by-step before producing its final sentiment label (Positive, Negative, or Neutral).
- Capture both the reasoning output and the final predicted label.

Expected Insight:

Students should demonstrate the ability to design effective prompts that elicit explicit reasoning from a large language model, interpret its reasoning quality critically, and compare these results with fine-tuned Transformer outputs.

Part 2: Comparative Analysis (14%)

- Compare BERT-base, FinBERT, and GPT-OSS predictions both quantitatively (macro-F1, accuracy) and qualitatively (example analysis).
- Identify at least five examples each of:
    - Correctly classified by FinBERT but misclassified by BERT.
    - Misclassified by both models.
- Analyse and explain these examples (e.g., negations, mixed tone, multi-clause headlines, neutral ambiguity).
- Present at least five to ten examples where GPT-OSS reasoning aligns or conflicts with FinBERT/BERT predictions.

Part 3: Reflection and Discussion (6%)

- Reflect on the following:
  - o The significance of domain-specific language models in financial NLP tasks.
  - o Risks of overfitting and biases in sentiment classification.
  - o Ethical considerations in automating financial decisions using AI systems.
  - o Key differences between training BERT/FinBERT and reasoning with GPT-OSS.
- About GPT-OSS Discuss:
  - o How reasoning-based prompting differs from fine-tuning.
  - o The strengths and weaknesses observed in GPT-OSS reasoning chains.
  - o Practical and ethical implications of using large generative models for financial text analysis.
- Summarise lessons learned and possible directions for future improvements.

Part 4: Presentation (20%)

Each group will present their findings from Parts 1–3.
The presentation should summarise the entire analysis, highlight insights, and demonstrate understanding of technical and ethical dimensions.

Guidelines:

- Each group will have a total of 25 minutes; 15–20 minutes for the presentation and 5–10 minutes for questions and answers. Groups must strictly adhere to the allotted time; presentations exceeding the limit will be stopped.
- Slides must be submitted as Microsoft PowerPoint files via Turnitin 24 hours before the scheduled presentation date/time.
- All group members must participate in presenting.
- Presentations should clearly communicate the experimental setup, results, key findings, and reflections with a good Clarity of understanding and content, Explanation of methods and experiments, Interpretation of results and reasoning discussion, Response to questions and depth of insight.
- **A presentation schedule and group list will be released on Moodle in Week 8.**
- Any exceptional circumstances must be reported to the lecturer **at least a week in advance of the scheduled presentation.**

In case of failure to submit this First Sit Assignment:

In any circumstances a student fails to submit their assignments in the first sit, a resit paper paper will be supplied later in which each student must work independently. Forming groups is not feasible due to the varied circumstances and availability of students during the resit period.

If circumstances change; for example, IF more than one student becomes available and IF grouping becomes feasible, this will be communicated before the resit assignment brief announcement on the Moodle. However, the overall style and structure of the assignment evaluation as below will remain unchanged. The only change that will be applied is that, instead of individual submissions, groups will

submit this assignment, and each student will be evaluated as part of their group in the similar manner as the first sit.

LATE SUBMISSIONS

Penalised according to UEL regulations below.

We strongly suggest that you try to submit all coursework by the deadline set as meeting deadlines is expected in employment. However, in our regulations, UEL has permitted students to be able to submit their coursework up to 24 hours after the deadline. The deadline will be published in your module guide. Coursework, which is submitted late, but within 24 hours of the deadline, will be assessed but subjected to a fixed penalty of 5% of the total marks available (as opposed to marks obtained).

However, you must be very careful when you are submitting your assessment. If you submit your work twice, once using the original deadline link and then again using the late submission link, your assignment will be graded as late with the 5% deduction.

Please note that if you submit twice, once before the deadline and once during the 24-hour late period, then the second submission will be marked and 5% deducted.