

Assignment # 3 – Topic Modeling and Sentiment Analysis in Financial Text

Instructor: Dr. Ahmad Raza Shahid

Course Information

Course: Deep Learning for Perception (CS4045)

Assignment: Design and Evaluation of Topic Modeling + Sentiment Analysis Pipelines

Due Date: As announced on Google Classroom (GCR)

Total Points: 100

Format: Jupyter/Colab Notebooks + Formal Report (PDF)

Assignment Philosophy

This assignment is **experimental and research-oriented**, focused on building a complete **financial text understanding pipeline**.

Students will design and evaluate several modern NLP components for financial **sentiment analysis**, including:

- Topic modeling (LDA)
- Transformer-based sentiment analyzers (FinBERT)
- Local LLM-based sentiment analysis
- Retrieval-Augmented Generation (RAG) for improved sentiment reasoning

You will compare these methods, justify your design decisions, and determine whether **fine-tuning** is required to achieve high performance.

Think like an experimental researcher—iterate, measure, and justify every step.

Dataset

- Download from Google Classroom (GCR).
-

Rules & Constraints (Strict)

- You must implement **three** sentiment analysis systems:
 1. **FinBERT-based Sentiment Analysis**
 2. **Local LLM-based Sentiment Analysis**
 3. **RAG-Enhanced Sentiment Analysis**

- You must perform **LDA topic modeling** before sentiment analysis.
 - You must implement RAG yourself (FAISS + retrieval + LLM).
 - No transformer fine-tuning unless the Fine-Tuning Rule triggers it.
 - No external RAG pipelines or classification services.
 - All results must be reproducible (set random seeds).
-

TASKS

1. Data Preprocessing & EDA

- Load your finance dataset
 - Clean and preprocess text (lowercase, punctuation removal, tokenization, stopword removal)
 - Provide basic dataset insights (class balance, sample sentences, length distribution)
-

2. Topic Modeling (LDA)

- Perform LDA topic modeling
 - Experiment with different topic numbers
 - Report:
 - Topic keywords
 - Topic coherence
 - Topic assignments for each sentence
 - Briefly interpret each topic (e.g., “market movement”, “company performance”, etc.)
-

3. Sentiment Analysis Using Three Approaches

A. FinBERT Sentiment Analysis

- Run FinBERT on your dataset
- Compute evaluation metrics:
 - Accuracy
 - Precision, Recall, F1
 - Confusion Matrix

B. Local LLM Sentiment Analysis

- Use any local or open-source LLM
- Perform zero-shot or few-shot sentiment analysis
- Compute the same metrics as FinBERT
- Compare performance

C. RAG-Enhanced Sentiment Analysis

- Create sentence embeddings (SentenceTransformers or similar)
- Build a **FAISS index**
- For each sentence:
 - Retrieve top-k similar sentences
 - Provide retrieved context to an LLM
 - Let the LLM produce a sentiment output
- Evaluate performance

4. Comparative Analysis

Analyze and compare:

Method	Key Points to Analyze
FinBERT	Strengths/weaknesses on financial text
Local LLM	Does it outperform FinBERT? Why?
RAG + LLM	Does retrieval help? When? Why or why not?

Include:

- Metric tables
- Confusion matrices
- Error case discussion

5. Fine-Tuning Rule (Mandatory Requirement)

After evaluating all three methods:

You must check whether fine-tuning is required:

- If **any** method achieves $\geq 90\%$ accuracy, then **fine-tuning is NOT required**.
- If **all** methods perform **below 90%**, then you **must fine-tune one model** (FinBERT or your LLM).

If Fine-Tuning Is Required:

- Fine-tune the chosen model on your dataset
- Recompute all metrics
- Achieve **at least 90% accuracy** after fine-tuning
- Document training details in your report

⚠ This step is used for grading.

Students who skip the fine-tuning decision or ignore required fine-tuning will lose marks.

Objective & Grading Policy

Objective: Design, implement, evaluate, and compare multiple sentiment analysis systems on financial text.

Grading (100 points):

Component	Points
Data Preprocessing & EDA	10
LDA Topic Modeling	20
FinBERT Sentiment Analysis	15
Local LLM Sentiment Analysis	15
RAG-Enhanced Sentiment Analysis	20
Fine-Tuning Decision & Implementation	10
Report Quality & Analysis	10
Total	100

Deliverables

You must submit:

1. Jupyter/Colab Notebooks

Must include:

- Preprocessing
- LDA topic modeling
- All three sentiment analysis systems
- Evaluation metrics
- Fine-tuning check + fine-tuning code (if needed)
- Code that reproduces your final accuracy

2. Formal Report (PDF, 4–8 pages)

Include:

- Abstract
- Dataset description
- Topic modeling results
- Model designs (FinBERT, LLM, RAG)
- Evaluation metrics & confusion matrices
- Fine-tuning justification
- Final analysis & insights

3. Model Checkpoints

- Fine-tuned model weights (if fine-tuned)
- Embeddings, FAISS index, and results CSV

4. README.txt

Instructions to run your notebooks and reproduce results.

Suggested Experimental Directions

- Try different numbers of LDA topics
- Try different embedding models for RAG
- Try different RAG retrieval depths ($k = 3, 5, 10$)
- Test different LLM prompting strategies
- Check whether RAG improves POS/NEG recall
- Try partial fine-tuning (last layers only)

- Try few-shot prompting for the LLM
 - Error analysis: manually inspect misclassified sentences
-

Evaluation & Reproducibility

- Instructors will re-run your evaluation cells
 - If your notebook does not reproduce your claimed accuracy, we use the reproduced accuracy
 - Set all seeds and ensure deterministic runs
-

Academic Integrity

- Ideas may be discussed, but code and reports must be your own
 - Copying code or results from external repositories without citation is not allowed
 - Plagiarism results in a zero
-

Tips for Success

- Build your pipeline step-by-step
 - Keep logs of experiments and hyperparameters
 - Run small tests before full experiments
 - Make RAG prompts clean and structured
 - Document everything clearly
-

Deadline & Submission

Submit a ZIP file named:

i22xxxx_name_A2.zip

Containing:

- **notebooks/**
- **report.pdf**

Upload via Google Classroom or the course portal.