

Assignment: Financial Data Extraction Using Open-Source LLMs

Objective: Your task is to use an open-source Large Language Model (LLM) trained specifically for handling financial data to extract key entities from the provided pdf documents.

Please access the pdfs here:

<https://drive.google.com/drive/folders/1BtpCkwql7LnMXtR2fjoaQUdqzVleIJGs?usp=sharing>

Use the selected LLM to extract the following financial entities from the pdfs:

- **Company Name**
- **Report Date**
- **Profit Before Tax**
- (Bonus: Extract additional relevant financial details if possible.)

Deliverables:

- **Code Implementation:** A well-documented Jupyter Notebook or Python script.
- **Expected Output Format:** A JSON file with a structured key-value format.