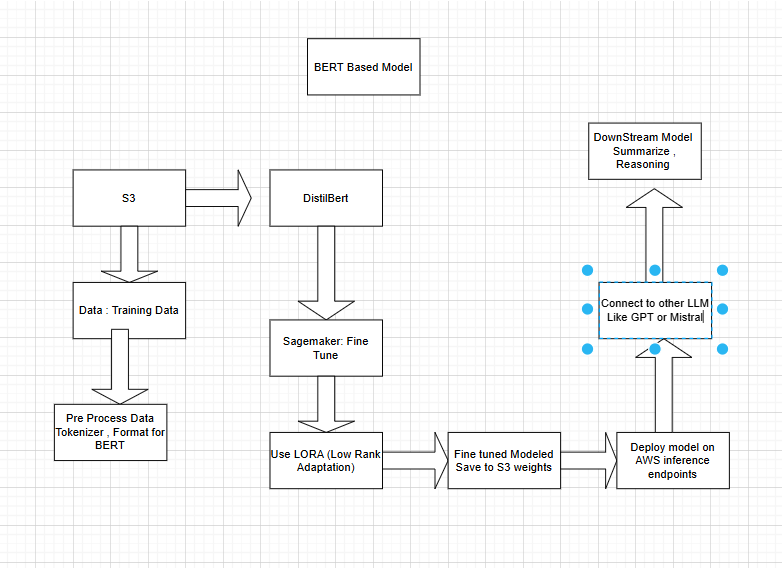
**CLI AI Agent for Industry‑Specific Event Identification in Email Threads**

Objective:

* Takes **raw email threads** as input
* Understands the **industry context** (e.g., finance, tech, logistics)
* Identifies and **highlights key events** relevant to that industry (e.g., "compliance deadline," "feature release," "shipment delay")



* How would I identify key events in an industry?

Perform Industry research , will search for standard operating events and KPI , use existing Taxonomies and business Glossaries

Sources

1. Public datasets (Enron etc)
2. Industry Documentation
3. Conversation with Domain Expert

* What data would I use to test your agent, and how would you gather and annotate it?

**Data Strategy:**

* Use a mix of **synthetic + semi-synthetic** email threads.

#### **a. Synthetic Data**

* Generate 500–1000 synthetic email threads per industry using GPT-4.

#### **b. Semi-Synthetic (based on real data)**

* Take public corpora like the **Enron Email Dataset**.
* Rewrite message contents to mimic key events across domains.

#### **c. Annotation Process**

* Use **Label Studio** or custom Python scripts to label:  
  + Event Type (from your ontology)
  + Trigger Sentence (or span)
  + Confidence (optional)

#### **d. Ensure Diversity**

* Cover various industries
* Include neutral (non-event) threads for false-positive testing
* Include long threads, short threads, nested replies

**What model(s) would you use? Would you use multiple models? Why or why not?**

BERT ( Encoder Based Models )

Fine Tune using LORA

Will use DistillBERT for classification and then use another LLM to get the response from BERT process it and then give to the end user

This way both of them are specialized and would perform their respective tasks well.  
BERT is an encoder based model : Good for understanding text and creating semantic meaning from vectors

Whereas GPT , Mistral is decoder only : Good for generating Text (so human speech )

**5. What technologies would you use?**

Model Training : Hugging Face Transformer + LORA

Cloud : AWS ( Sagemaker , S3 , EC2, endpoints , Lambda , API Gateway

CLI : argparse

Python and its related libraries for preprocessing

**Explain the workflow in detail.**

#### **Step 1: CLI Input**

**Accepts:  
  
 bash  
  
python cli\_agent.py --file thread.txt --industry finance --export results.json**

#### **Step 2: Preprocessing**

* **Strip signatures, greetings, disclaimers**
* **Split thread by message and tokenize text**
* **Upload processed version to S3 for logging (optional)**

#### **Step 3: Load Model from S3**

**Automatically loads appropriate fine-tuned BERT-LoRA model for specified industry  
  
 python  
  
model\_path = f"s3://models/bert-lora-{industry}.pt"**

#### **Step 4: Prediction**

* **Tokenize input using model tokenizer**
* **Run forward pass → extract predicted events with type + location**

**Return structured output:  
  
 json  
  
[**

**{**

**"event": "Regulatory Update",**

**"line": "We must comply with new SEBI guidelines this quarter.",**

**"confidence": 0.89**

**}**

**]**

#### **Step 5: LLM Post-Processing**

* **If enabled:**
  + **Summarize events**
  + **Generate automatic replies**
  + **Add contextual reasoning or action items**

#### **Step 6: Output**

* **Display in CLI with Rich formatting**
* **Optionally export to JSON, markdown, or HTML for further use**