**End to End Model Structure**

**Primary Goal** : To make a chatbot fine-tuned for defense organizations Q/A purposes. The model should be light weight and high on accuracy. (decent accuracy percent already achieved as Rana sir said).

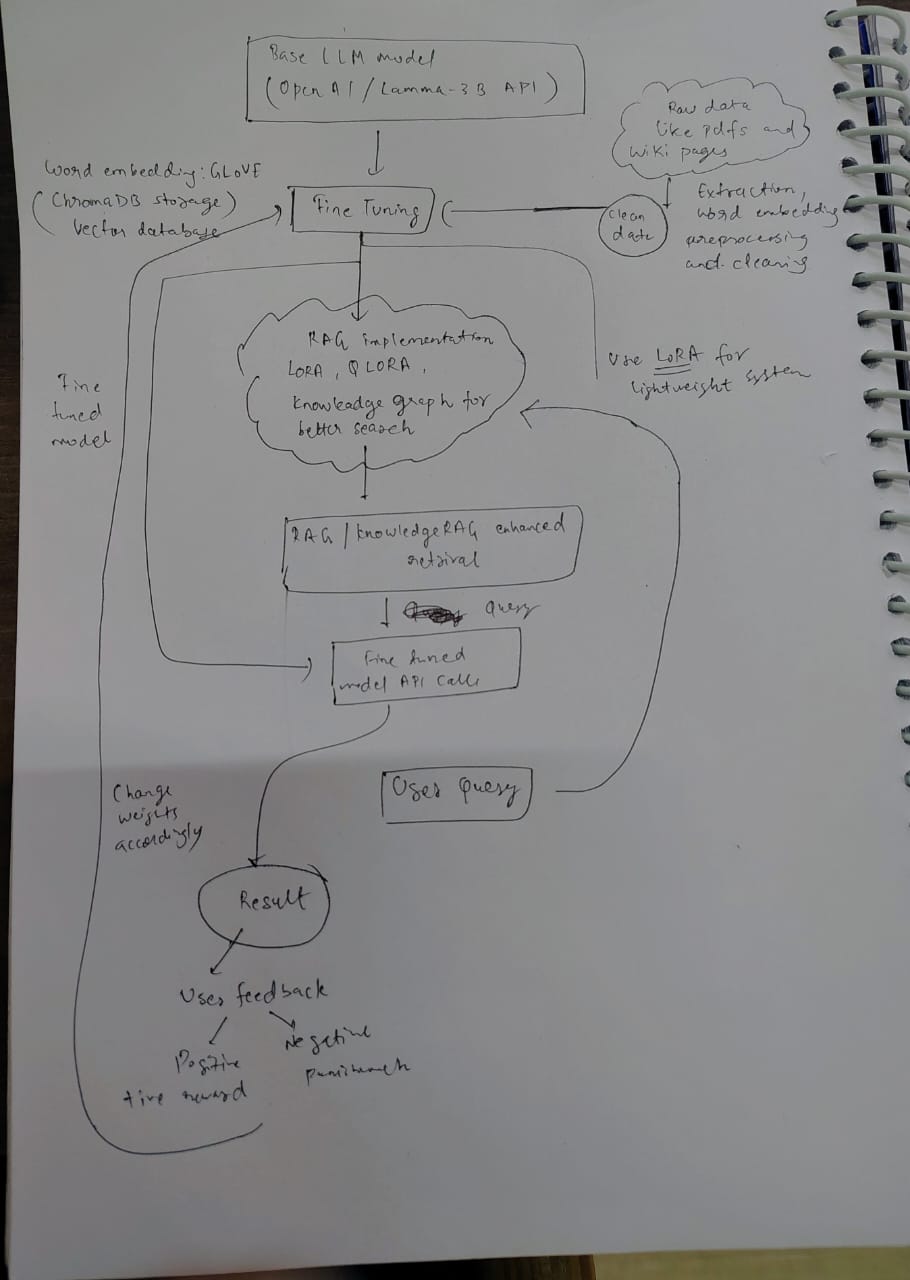
Thought Process:

Take a LLM API like GPT-4 or Llama-3B. Fine-tune it using techniques like LoRA or QLoRA (more techniques available I guess, research required). For training we have key value dictionary.

Apply preprocessing , use large Word embedding models like GloVE ( this also need to be researched, what is best ) for words, store them in vector storage like chroma DB etc. Implement RAG and Knowledge graph for efficient retrieval, Train the model.

Now for feedback we want a RL chain at the backend. (research required)

Workflow chart :



Models :

DeepSeek R1 – Good for reasoning and maths as they say, I don’t think so it is useful in our case, as we focus on accurate Q/A. –can change transformer architecture at some checkpoints.

Llama – 3B – can do, already done by rana sir

OpenAI- 4o (Paid) : no transformer changes, but can use training based API calls for fine tuning

Mistral – 7B : We can try. Beats Llama on almost all tasks. Also light weight and can run on mac M3 I think

I Think we have to try all these and check model performance accordingly, we cant choose one directly and work on that

Vector Database – Chroma DB , to store vector embeddings.

**FINE TUNING!**

<https://towardsdatascience.com/fine-tuning-large-language-models-llms-23473d763b91/?sk=fd31e7444cf8f3070d9a843a8218ddad>

Learned about LoRA and QLoRA for fine tuning the base model.

Saw how to change the positional encodings, feed forward layers manually using pytorch and huggingface

**Reinforcement Learning**

**Learned** about RLHF, we have to manually score the chatbot responses , and then train again by (key,value, score) pairs.

**PPO algorithm**

Mathematically, PPO optimizes the policy using a "Clipped Objective":

A black screen with white text

Description automatically generated

**Initial Implementation Plan**

1. **Model Evaluation:** Compare performance of shortlisted models (DeepSeek R1, Llama-3B, GPT-4o, Mistral-7B) on Q&A tasks.
2. **Fine-Tuning:** Use LoRA/QLoRA on the selected model with a key-value dictionary.
3. **Vector Storage:** Implement ChromaDB for efficient embedding retrieval. (RAG)
4. **Reinforcement Learning Integration:** Research and implement an RL-based feedback mechanism.
5. **Testing & Validation:** Evaluate the model’s accuracy and optimize further.