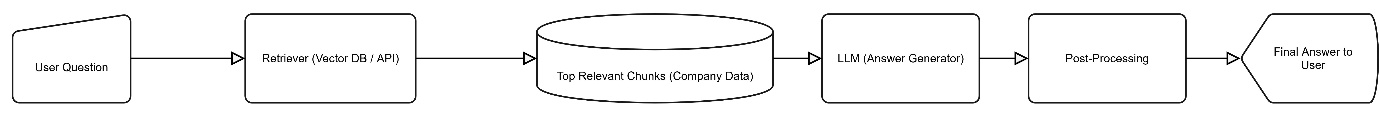
Task 3



First of all, I would like to apologize because I am still new and not yet fully understand about Large Language Models (LLM). However, I have tried my best to study the concept and summarize how a company could build an internal chat assistant that works like ChatGPT but uses only internal databases.

* **Data Preparation**
  + Collect internal company documents and databases.
  + Clean and format the data, then split into smaller sections (chunks).
  + Store these chunks in a **vector database** (or make them accessible through internal APIs).
* **Search / Retrieval**
  + When a user asks a question, the system first looks up the most relevant chunks in the vector database (or calls the company API) to ensures the LLM only sees company-approved data.
* **Answer Generation (LLM)**
  + The retrieved chunks are passed into the LLM (like ChatGPT or an open-source model).
  + The LLM generates a natural language answer **based only on those chunks.**
* **Preventing Hallucination**
  + If no relevant chunk is found, the assistant should reply “I don’t know based on my data.”
  + For sensitive facts (e.g., employee info, transaction history), verify directly with the internal API instead of free text generation.
* **Evaluation & Improvement**
  + Build a test set of typical employee questions + correct answers.
  + Measure accuracy, relevance, and how often the model hallucinates.
  + Collect user feedback and update data + pipeline regularly.

Task 4

To build an end-to-end data platform I will use Alibaba services as example. The first step is data ingestion. Internal and external data are collected through streaming jobs (e.g., Apache Flink) or batch jobs (via SFTP, function compute, or orchestrators like Airflow). The ingested data lands in a data lakehouse environment (Alibaba OSS, Apache Paimon, MaxCompute) where it is stored, cleaned, and transformed. Once the data is standardized, it is organized with metadata and access control. From there, the clean data can be used for downstream needs such as AI/ML projects, reporting dashboards, or advanced data analysis.

To deploy this system in cloud platforms such as GCP, AWS, or Azure, the same concept applies: ingestion services) feed into datalake or datalakehouse (Depends on requirement) with orchestration handled by managed pipelines (Airflow, Dataflow, Glue). For model evaluation and analytics, the curated data is connected to ML platforms where models are trained, validated, and monitored. TO endure countinuous updates, we can use CI/CD pipelines and apply governance and monitoring (data catalog, and access management) to prevent quality issues and ensure security.

Task 5

The future of fintech in investment banking is expected to be influenced by automation, data-driven decision-making, and increased democratization of access to financial markets. Historically, investment banking has been primarily controlled by large institutions with substantial human resources, but fintech is progressively altering this landscape. Technologies such as artificial intelligence and machine learning are already being utilized in areas like risk modelling, portfolio optimization, and deal sourcing. As these technologies advance, they will lessen dependence on manual processes, speed up transaction cycles, and enhance the accuracy of market trend forecasts. Consequently, bankers will allocate less time to repetitive tasks and focus more on strategic advisory functions.

Another significant trend is the emergence of decentralized finance (DeFi) and blockchain technologies. For investment banking, this could signify a transformative shift in the operation of capital markets, leading to more transparent and efficient systems for clearing, settlement, and fundraising. The tokenization of assets may enable fractional ownership of traditionally illiquid investments, such as private equity, real estate, or infrastructure projects. This development paves the way for a wider investor base and introduces innovative financing models that could redefine underwriting, syndication, and secondary market trading.