

# CSE 676-B Deep Learning

## Final Project

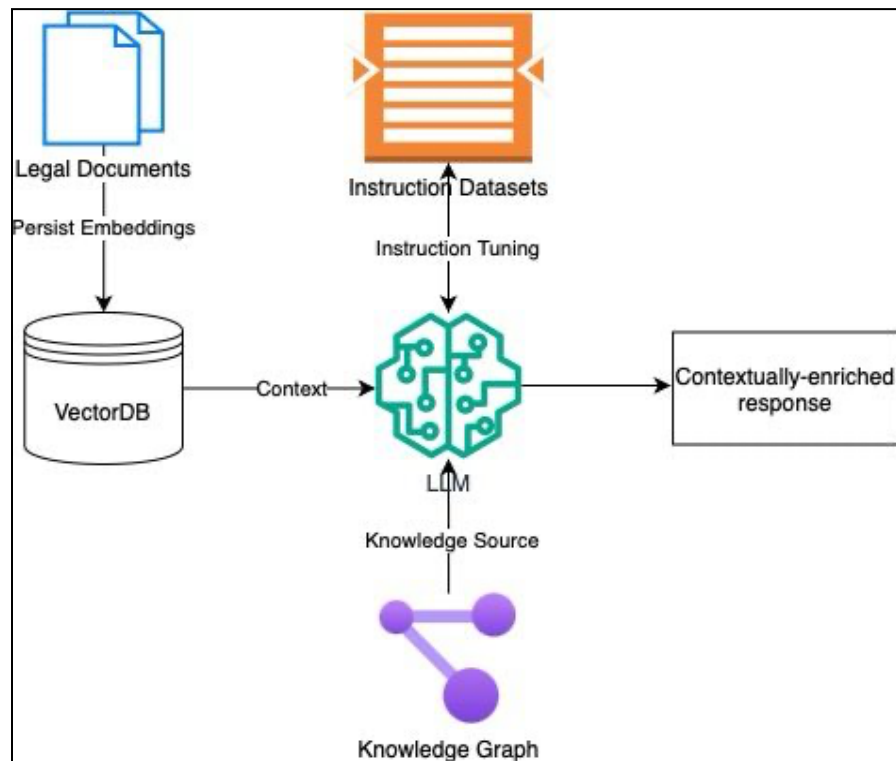
### Project Checkpoint

## Contextually-Aware LLMs Using Knowledge Graphs

### Objective :

We are planning to build a Question-Answering platform for the legislative domain by combining **Retrieval-Augmented Generation (RAG)** and **Knowledge Graphs** to provide contextually rich, accurate answers to legal questions. We aim to enable intelligent interaction with large collections of legal documents and provide detailed responses grounded in factual knowledge. Our main objective is to create a system that is lightweight and accessible. For this, we will leverage multiple **small language models** within the RAG pipeline, augmented by knowledge graphs, to achieve accuracy comparable to state-of-the-art large-scale LLMs. This will enable deployment on **edge devices and low-resource environments**, extending usability beyond centralized servers or cloud-based setups.

### Architecture :



## Progress So Far:

### 1. RAG pipeline implementation:

- a. We implemented a standard RAG pipeline using three small models relevant to QA tasks using the LangChain framework:
  - i. google/flan-t5-base
  - ii. declare-lab/flan-alpaca-base
  - iii. allenai/unifiedqa-t5-base
- b. We used **SentenceTransformers (all-MiniLM-L6-v2)** for embedding generation.
- c. We used **LangChain's RecursiveCharacterTextSplitter** to chunk legal documents into manageable segments.
- d. We used **FAISS** as the vector store for efficient document retrieval.
- e. Created a Prompt with relevant contexts from the FAISS vector store.

### 2. Evaluation Metrics:

We evaluated responses generated by the baseline models via a basic RAG pipeline using the SQuAD evaluation metrics - Exact Match and F1 score, which allowed us to benchmark our system's response quality against standard QA datasets.

### 3. FAISS Vector DB Setup

- a. We created a vector database of legal document chunks using a scalable embedding process and enabled persistent storage and reuse of the FAISS index for query answering.
- b. In our system, we are :
  - i. Reading all documents from a folder.
  - ii. Splitting them into chunks.
  - iii. Embedding them into vector form.
  - iv. And finally, storing them in a FAISS database for fast similarity-based retrieval.

### 4. QA Baseline model :

- a. First, we load the FAISS index.
- b. We retrieve the top 4 relevant chunks using Maximal Marginal Relevance (MMR).
- c. Then, format the context into a QA-friendly prompt.
- d. Finally, we use a small text2text model (future steps: Using a legal-aware model) to generate a concise, context-grounded answer.
- e. This will work well for our use case as legal documents are often long and complex, and chunking + retrieval helps isolate the relevant portions, and the prompting structure keeps the output factual, concise, and user-friendly.

## Further Implementation Plan :

### 1. Use Case implementation : Legislative Domain.

- a. Our final QA system will be specifically designed for legal professionals, researchers, and law students and will be tailored for law-related questions.
- b. Legal documents like case laws, regulations, and statutes will be ingested, chunked, embedded, and indexed so that when the user inputs legal queries, the system will respond with contextually enriched answers.

### 2. Knowledge Graph integration :

We will integrate a Custom-built graph Knowledge-Graph using Neo4j for specific laws, sections, and precedents to:

- i. Link legal entities and cases.
- ii. Provide structured context to augment generative answers.
- iii. Improve fact-grounding and traceability in legal answers.

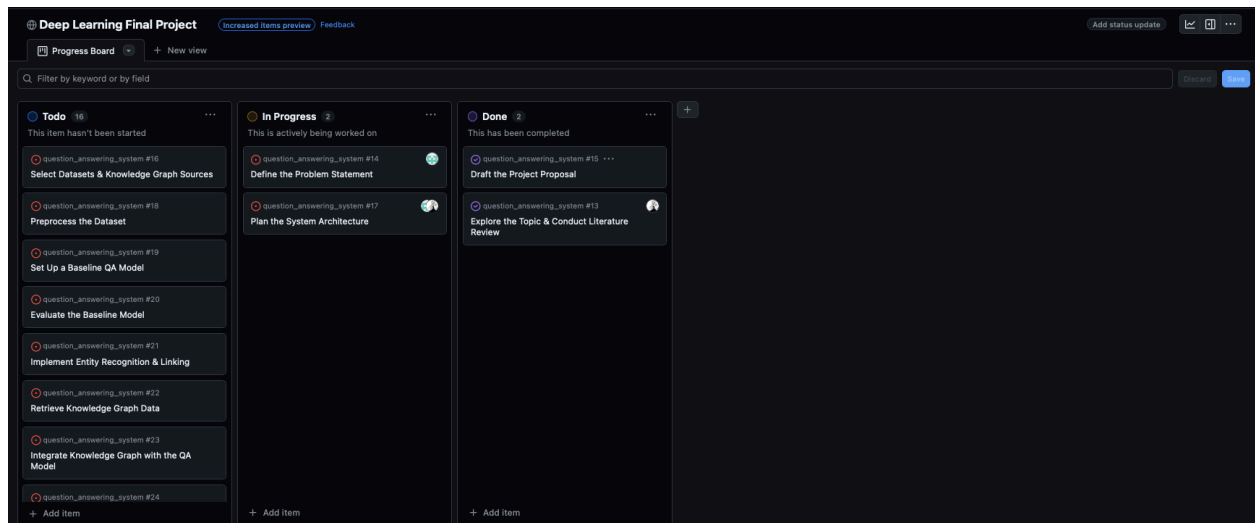
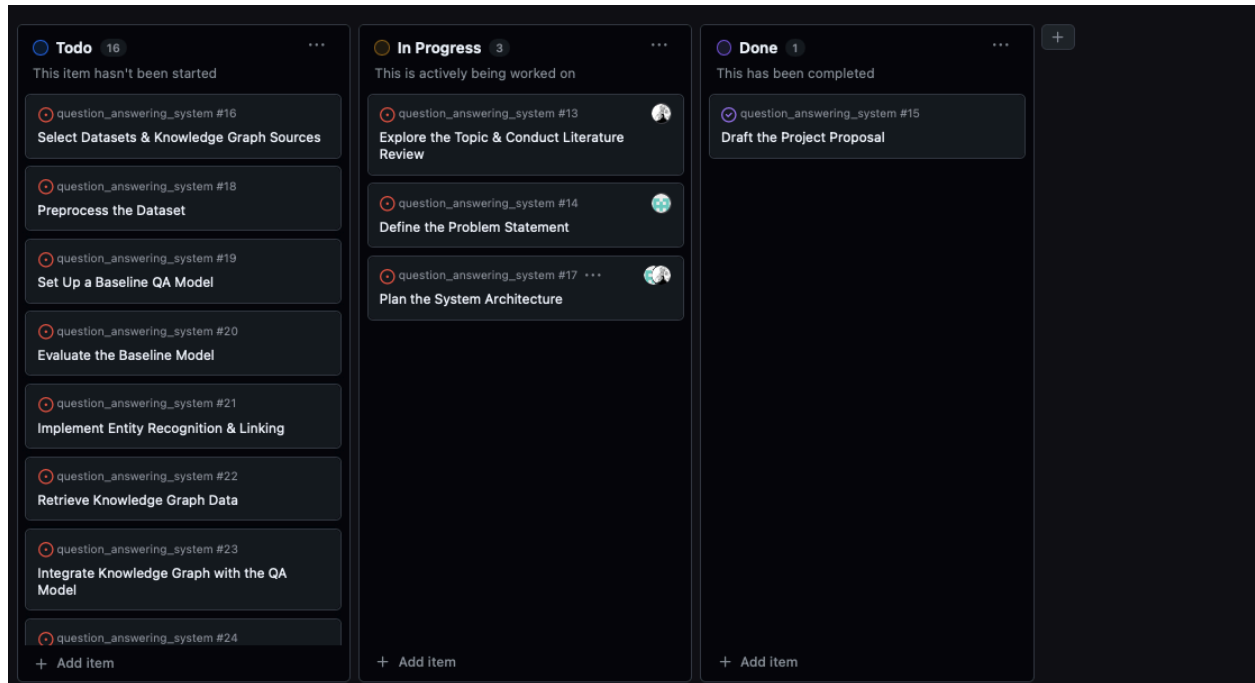
### 3. Final Architecture will include :

- a. **Document Ingestion & Chunking:** Using LangChain's loaders and splitters  
We will load raw legal texts from different sources and use character-based chunking to split long texts.
- b. **Embedding & Retrieval:** SentenceTransformer + FAISS  
We will use SentenceTransformer to convert chunks to dense vectors and retrieve similar chunks according to the user's legal question by storing all document chunks and allowing efficient semantic retrieval with FAISS.
- c. **Answer Generation:**
  - i. We will use Small transformer-based LLMs (listed above) to generate a coherent answer from the retrieved contexts and Knowledge Graph.
  - ii. After this, we will instruct tune our models with legal information to give more accurate results.
- d. **Knowledge Graph Augmentation:**
  - i. We will identify Legal entities and the relationships between those entities via NER + linking.
  - ii. During answer generation, contextual facts and additional knowledge will be retrieved from the KG to improve answer precision and reduce hallucinations.

### 4. Evaluation Metrics :

- a. We will include custom legal domain RAG benchmarks to examine retriever performance and downstream question-answering performance.
- b. And will work on hallucination reduction with KG support.
- c. LegalBench - <https://hazyresearch.stanford.edu/legalbench/getting-started/>

## Project Board Screenshots



Aayush Subramaniam - 50605671  
Meghna Verma - 50604872

**Progress Board** + New view

Q. Filter by keyword or by field

**Todo** 15  
This item hasn't been started

- question\_answering\_system #27  
Conduct Error Analysis
- question\_answering\_system #28  
Run Extensive Model Evaluation
- question\_answering\_system #29  
Perform Case Studies & User Testing
- question\_answering\_system #30  
Prepare a Report & Documentation
- question\_answering\_system #31 ...  
Prepare for the Presentation
- question\_answering\_system #32  
Submit Final Project & Present Findings
- question\_answering\_system #33  
QA RAG Pipeline using Langchain and FAISS DB

+ Add Item

**In Progress** 3  
This is actively being worked on

- question\_answering\_system #17  
Plan the System Architecture
- question\_answering\_system #19  
Set Up a Baseline QA Model
- question\_answering\_system #20  
Evaluate the Baseline Model

+ Add Item

**Done** 3  
This has been completed

- question\_answering\_system #15 ...  
Draft the Project Proposal
- question\_answering\_system #13  
Explore the Topic & Conduct Literature Review
- question\_answering\_system #14  
Define the Problem Statement

+ Add Item

**Deep Learning Final Project** (Increased items preview) Feedback

**Progress Board** + New view

Q. Filter by keyword or by field

**Todo** 14  
This item hasn't been started

- Implement Multi-Hop Reasoning
- question\_answering\_system #26  
Improve Answer Generation
- question\_answering\_system #27  
Conduct Error Analysis
- question\_answering\_system #28  
Run Extensive Model Evaluation
- question\_answering\_system #29  
Perform Case Studies & User Testing
- question\_answering\_system #30  
Prepare a Report & Documentation
- question\_answering\_system #31  
Prepare for the Presentation
- question\_answering\_system #32  
Submit Final Project & Present Findings

+ Add Item

**In Progress** 4  
This is actively being worked on

- question\_answering\_system #17  
Plan the System Architecture
- question\_answering\_system #19  
Set Up a Baseline QA Model
- question\_answering\_system #20  
Evaluate the Baseline Model
- question\_answering\_system #33 ...  
QA RAG Pipeline using Langchain and FAISS DB

+ Add Item

**Done** 3  
This has been completed

- question\_answering\_system #15 ...  
Draft the Project Proposal
- question\_answering\_system #13  
Explore the Topic & Conduct Literature Review
- question\_answering\_system #14  
Define the Problem Statement

+ Add Item

**Deep Learning Final Project** (Increased items preview) Feedback

**Progress Board** + New view

Q. Filter by keyword or by field

**Todo** 14  
This item hasn't been started

- Implement Multi-Hop Reasoning
- question\_answering\_system #26  
Improve Answer Generation
- question\_answering\_system #27  
Conduct Error Analysis
- question\_answering\_system #28  
Run Extensive Model Evaluation
- question\_answering\_system #29  
Perform Case Studies & User Testing
- question\_answering\_system #30  
Prepare a Report & Documentation
- question\_answering\_system #31  
Prepare for the Presentation
- question\_answering\_system #32  
Submit Final Project & Present Findings

+ Add Item

**In Progress** 3  
This is actively being worked on

- question\_answering\_system #19  
Set Up a Baseline QA Model
- question\_answering\_system #20  
Evaluate the Baseline Model
- question\_answering\_system #33 ...  
QA RAG Pipeline using Langchain and FAISS DB

+ Add Item

**Done** 4  
This has been completed

- question\_answering\_system #15 ...  
Draft the Project Proposal
- question\_answering\_system #13  
Explore the Topic & Conduct Literature Review
- question\_answering\_system #14  
Define the Problem Statement
- question\_answering\_system #17  
Plan the System Architecture

+ Add Item

Aayush Subramaniam - 50605671  
Meghna Verma - 50604872

**Progress Board** + New view

Q Filter by keyword or by field

**Todo** 13  
This item hasn't been started

- Implement Multi-Hop Reasoning
- question\_answering\_system #26 Improve Answer Generation
- question\_answering\_system #27 Conduct Error Analysis
- question\_answering\_system #28 Run Extensive Model Evaluation
- question\_answering\_system #29 Perform Case Studies & User Testing
- question\_answering\_system #30 Prepare a Report & Documentation
- question\_answering\_system #31 Prepare for the Presentation
- question\_answering\_system #32 Submit Final Project & Present Findings

+ Add Item

**In Progress** 5  
This is actively being worked on

- question\_answering\_system #19 Set Up a Baseline QA Model
- question\_answering\_system #20 Evaluate the Baseline Model
- question\_answering\_system #33 QA RAG Pipeline using Langchain and FAISS DB
- question\_answering\_system #16 Select Datasets & Knowledge Graph Sources
- question\_answering\_system #34 ... Checkpoint Submission

+ Add Item

**Done** 4  
This has been completed

- question\_answering\_system #15 ... Draft the Project Proposal
- question\_answering\_system #13 Explore the Topic & Conduct Literature Review
- question\_answering\_system #14 Define the Problem Statement
- question\_answering\_system #17 Plan the System Architecture

+ Add Item

**Deep Learning Final Project** Increased items preview Feedback

**Progress Board** + New view

Q Filter by keyword or by field

**Todo** 13  
This item hasn't been started

- Implement Multi-Hop Reasoning
- question\_answering\_system #26 Improve Answer Generation
- question\_answering\_system #27 Conduct Error Analysis
- question\_answering\_system #28 Run Extensive Model Evaluation
- question\_answering\_system #29 Perform Case Studies & User Testing
- question\_answering\_system #30 Prepare a Report & Documentation
- question\_answering\_system #31 Prepare for the Presentation
- question\_answering\_system #32 Submit Final Project & Present Findings

+ Add Item

**In Progress** 4  
This is actively being worked on

- question\_answering\_system #20 Evaluate the Baseline Model
- question\_answering\_system #33 QA RAG Pipeline using Langchain and FAISS DB
- question\_answering\_system #16 Select Datasets & Knowledge Graph Sources
- question\_answering\_system #34 ... Checkpoint Submission

+ Add Item

**Done** 5  
This has been completed

- question\_answering\_system #15 ... Draft the Project Proposal
- question\_answering\_system #13 Explore the Topic & Conduct Literature Review
- question\_answering\_system #14 Define the Problem Statement
- question\_answering\_system #17 Plan the System Architecture
- question\_answering\_system #19 Set Up a Baseline QA Model

+ Add Item

**Deep Learning Final Project** Increased items preview Feedback

**Progress Board** + New view

Q Filter by keyword or by field

**Todo** 12  
This item hasn't been started

- question\_answering\_system #21 Implement Entity Recognition & Linking
- question\_answering\_system #22 Retrieve Knowledge Graph Data
- question\_answering\_system #23 Integrate Knowledge Graph with the QA Model
- question\_answering\_system #24 Fine-tune the Knowledge-Enhanced Model
- question\_answering\_system #25 Implement Multi-Hop Reasoning
- question\_answering\_system #26 Improve Answer Generation
- question\_answering\_system #27 Conduct Error Analysis
- question\_answering\_system #28

+ Add Item

**In Progress** 5  
This is actively being worked on

- question\_answering\_system #20 Evaluate the Baseline Model
- question\_answering\_system #33 QA RAG Pipeline using Langchain and FAISS DB
- question\_answering\_system #16 Select Datasets & Knowledge Graph Sources
- question\_answering\_system #34 ... Checkpoint Submission
- question\_answering\_system #18 ... Preprocess the Dataset for Instruction Tuning

+ Add Item

**Done** 5  
This has been completed

- question\_answering\_system #15 ... Draft the Project Proposal
- question\_answering\_system #13 Explore the Topic & Conduct Literature Review
- question\_answering\_system #14 Define the Problem Statement
- question\_answering\_system #17 Plan the System Architecture
- question\_answering\_system #19 Set Up a Baseline QA Model

+ Add Item

Aayush Subramaniam - 50605671  
Meghna Verma - 50604872

**Deep Learning Final Project** [Increased items preview](#) [Feedback](#) [Add](#)

**Deep Learning Final Project** [Increased items preview](#) [Feedback](#) [Add status update](#)

**Progress Board** [+ New view](#)

Filter by keyword or by field

**Todo** 11  
This item hasn't been started

question\_answering\_system #24

Fine-tune the Knowledge-Enhanced Model

question\_answering\_system #25

Implement Multi-Hop Reasoning

question\_answering\_system #27

Conduct Error Analysis

question\_answering\_system #28

Run Extensive Model Evaluation

question\_answering\_system #29

Perform Case Studies & User Testing

question\_answering\_system #30

Prepare a Report & Documentation

question\_answering\_system #31

Prepare for the Presentation

question\_answering\_system #32

[+ Add item](#)

**In Progress** 6  
This is actively being worked on

question\_answering\_system #20

Evaluate the Baseline Model

question\_answering\_system #33

QA RAG Pipeline using Langchain and FAISS DB

question\_answering\_system #16

Select Datasets & Knowledge Graph Sources

question\_answering\_system #34

Checkpoint Submission

question\_answering\_system #18

Preprocess the Dataset for Instruction Tuning

question\_answering\_system #26

Improve Answer Generation [Prompt Engineering]

[+ Add item](#)

**Done** 5  
This has been completed

question\_answering\_system #15

Draft the Project Proposal

question\_answering\_system #13

Explore the Topic & Conduct Literature Review

question\_answering\_system #14

Define the Problem Statement

question\_answering\_system #17

Plan the System Architecture

question\_answering\_system #19

Set Up a Baseline QA Model

[+ Add item](#)

**Deep Learning Final Project** [Increased items preview](#) [Feedback](#)

**Progress Board** [+ New view](#)

Filter by keyword or by field

**Todo** 11  
This item hasn't been started

question\_answering\_system #21

Implement Entity Recognition & Linking

question\_answering\_system #22

Retrieve Knowledge Graph Data

question\_answering\_system #23

Integrate Knowledge Graph with the QA Model

question\_answering\_system #24

Fine-tune the Knowledge-Enhanced Model

question\_answering\_system #25

Implement Multi-Hop Reasoning

question\_answering\_system #27

Conduct Error Analysis

question\_answering\_system #28

Run Extensive Model Evaluation

question\_answering\_system #29

Perform Case Studies & User Testing

question\_answering\_system #30

Prepare a Report & Documentation

question\_answering\_system #31

Prepare for the Presentation

question\_answering\_system #32

[+ Add item](#)

**In Progress** 3  
This is actively being worked on

question\_answering\_system #16

Select Datasets & Knowledge Graph Sources

question\_answering\_system #18

Preprocess the Dataset for Instruction Tuning

question\_answering\_system #26

Improve Answer Generation [Prompt Engineering]

[+ Add item](#)

**Done** 8  
This has been completed

question\_answering\_system #15

Draft the Project Proposal

question\_answering\_system #13

Explore the Topic & Conduct Literature Review

question\_answering\_system #14

Define the Problem Statement

question\_answering\_system #17

Plan the System Architecture

question\_answering\_system #19

Set Up a Baseline QA Model

question\_answering\_system #34

Checkpoint Submission

question\_answering\_system #20

Evaluate the Baseline Model

question\_answering\_system #33

QA RAG Pipeline using Langchain and FAISS DB

[+ Add item](#)

Project Board Link - <https://github.com/users/Meghna0327/projects/1>