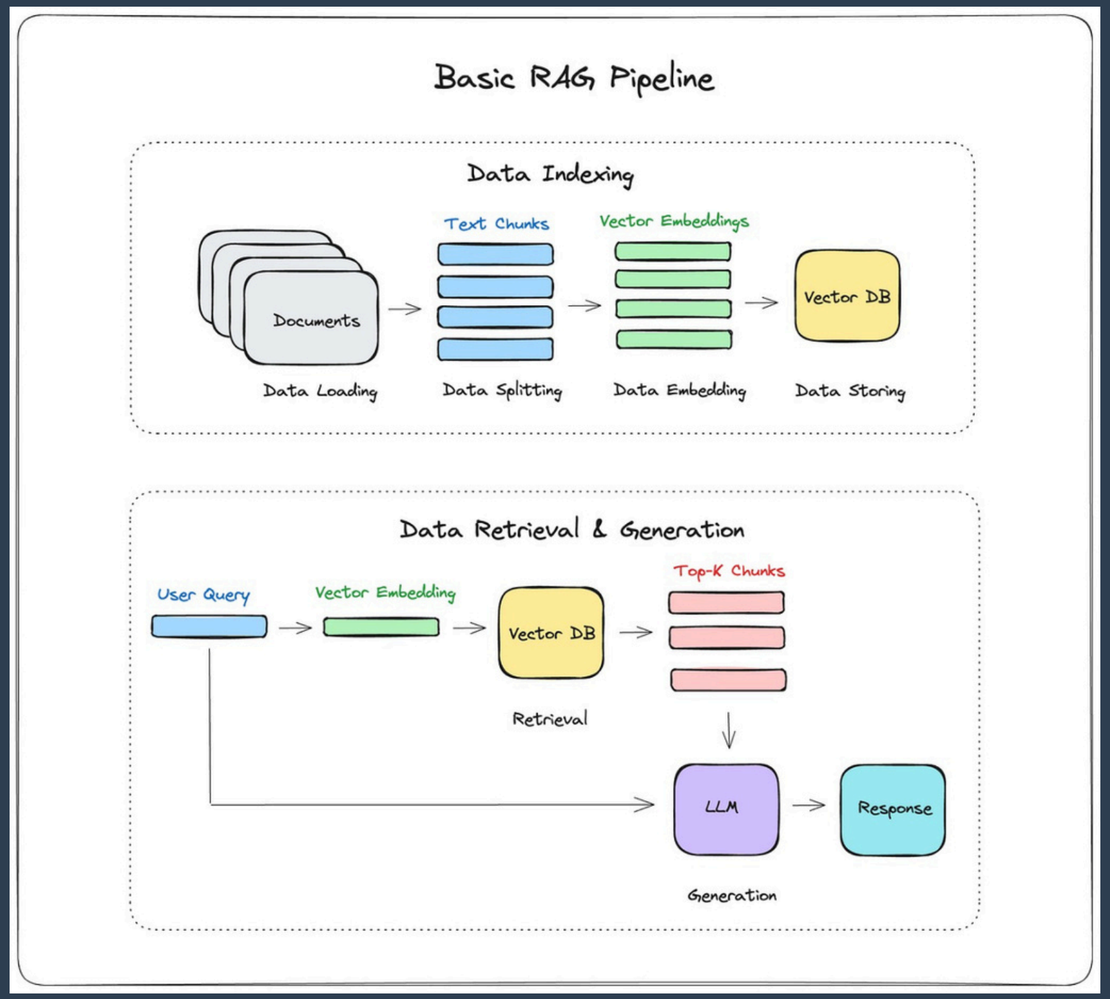


Introducing RAG+: Revolutionizing Data Access

Presented by G448
Mentor : Ashok Sharma Sir

Key Features

Context-Aware Semantic Retrieval

Integrates Top-K semantic similarity scoring to identify relevant documents even under vocabulary or phrasing mismatches, improving the contextual depth of responses.

Multi-Point Mapping

Implements many-to-many mapping between Application Points and Knowledge Points, allowing rich interlinking between legal issues, statutes, and case references.

Smart Summarization

Uses a Brief Summary Mode that distills long legal analyses into short, structured overviews highlighting:

- Key Issues
- Legal Principles
- Conclusions

Adaptive Analytics Engine

Monitors analytical performance metrics such as:

- Number of Statutes and Case Laws Utilized
- Total Words Generated
- Average Response and Retrieval Times

Configurable Output Control

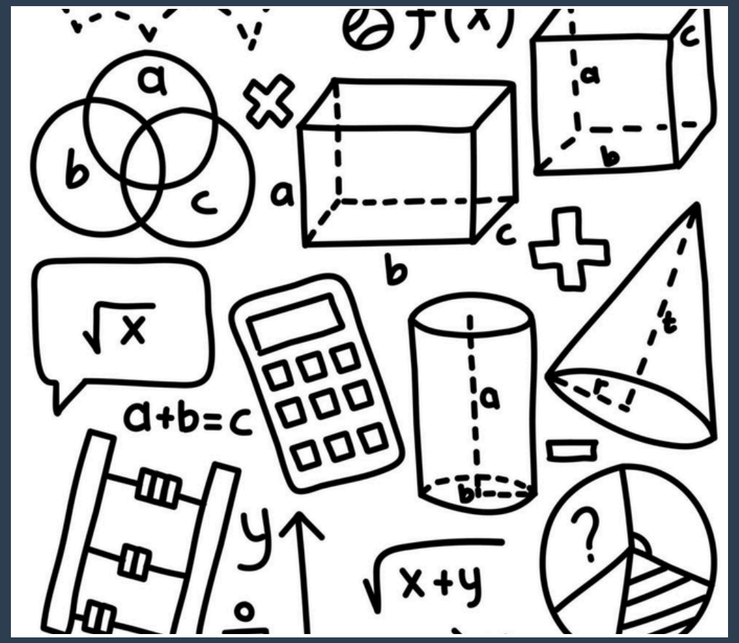
Allows customization of output parameters such as target word count, retrieval depth, and analysis scope, giving users flexibility over content length and detail.

Interactive Performance Dashboard

Displays real-time system statistics including:

- Query Count and Word Count
- Model in Use
- Processing Time
- Corpus and System Health Indicators

RAG+ Use Cases Explored

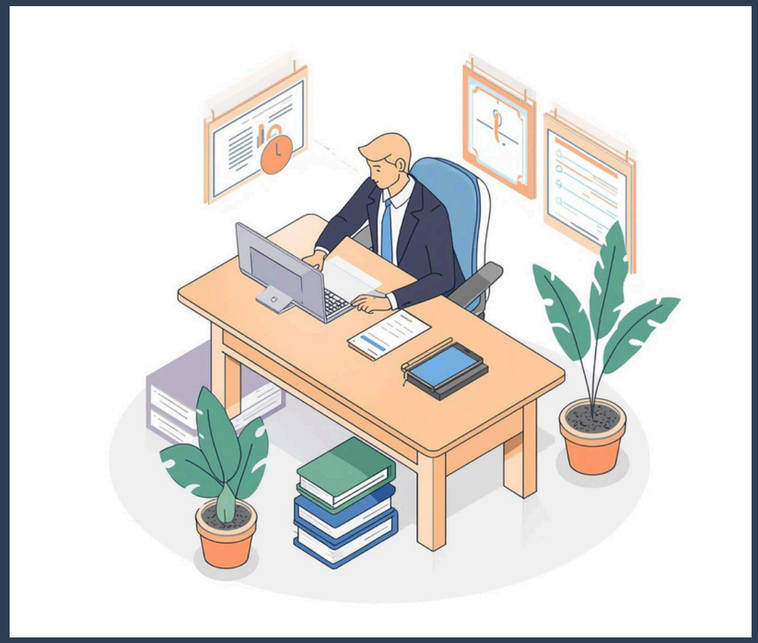


Mathematical

Solve complex Mathematical equation with high accuracy.

Healthcare

Summarize patient records quickly and effectively.



Legal

Review contracts efficiently with intelligent insights.

PROJECT TECH STACK

Frontend & UI

Streamlit
Web application framework for interactive UI

HTML/CSS
Custom styling with animations and responsive design

Javascript (embedded)
Javascript
Interactive elements and animations

AI & Machine Learning

Google Gemini 2.0
Large Language Model for text generation

SentenceTransforrers
Text embeding model (all-MiniLM-LG-v2)

Hugging Face Transformes
ML model ecosytem

Vector Database & Search

Pinecone
Cloud-based vector database for simiarity search

Semantic simiarity computior search

Cosine Silarity
Semantic simiarity computation

Improvements upon Base-Model

Dynamic Data Mapping:

Designed and implemented a flexible data model supporting many-to-many mapping between Application Points and Knowledge Points.
This enables richer contextual understanding, better reusability of knowledge components, and fine-grained control over content relationships during retrieval and generation.

Dynamic Application Point Creation:

Implemented an LLM-powered API pipeline that automatically generates new Application Points when none are retrieved from the existing database.
This allows the system to expand dynamically with minimal manual intervention, making it adaptive, self-evolving, and scalable for emerging knowledge domains.

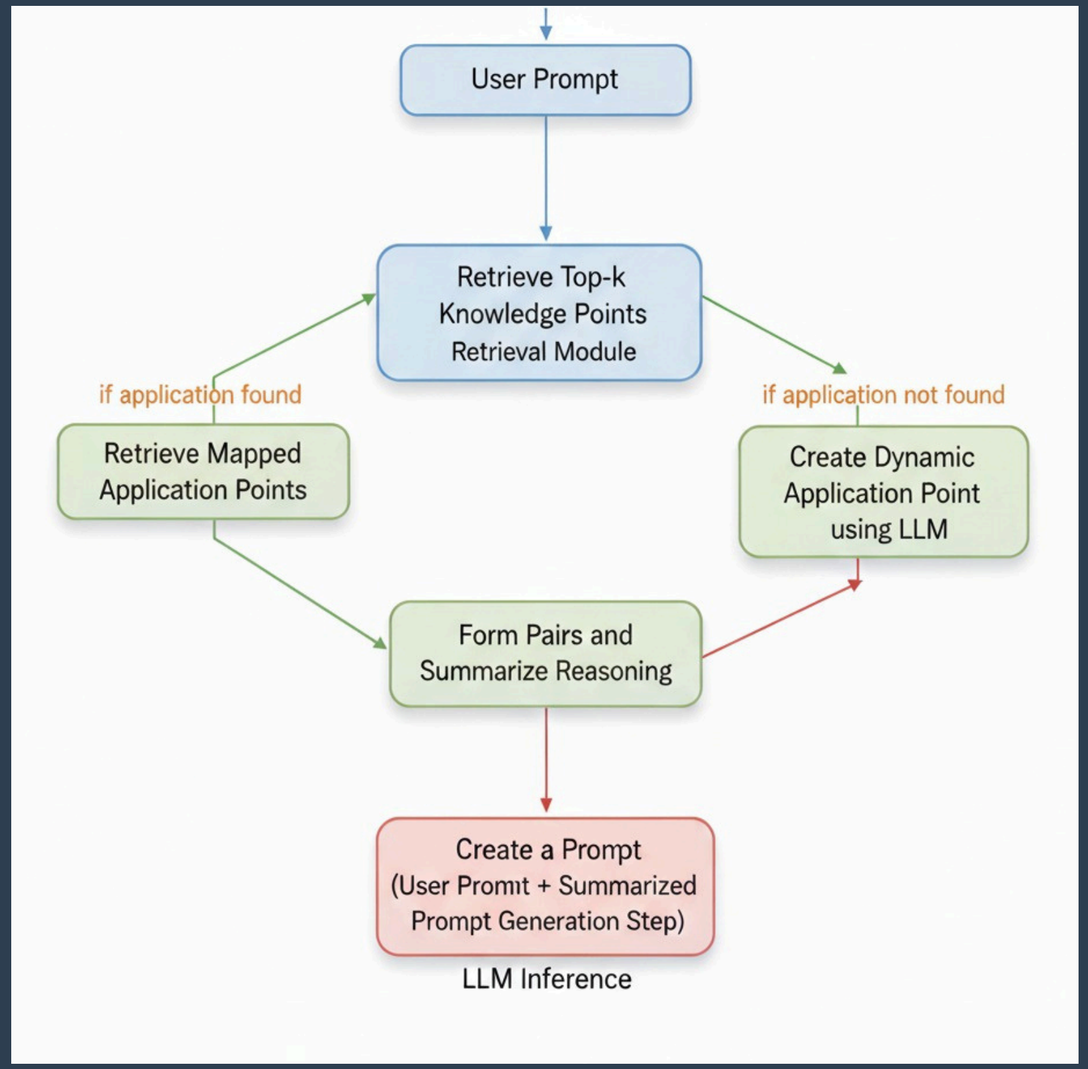
Custom Legal Dataset Development:

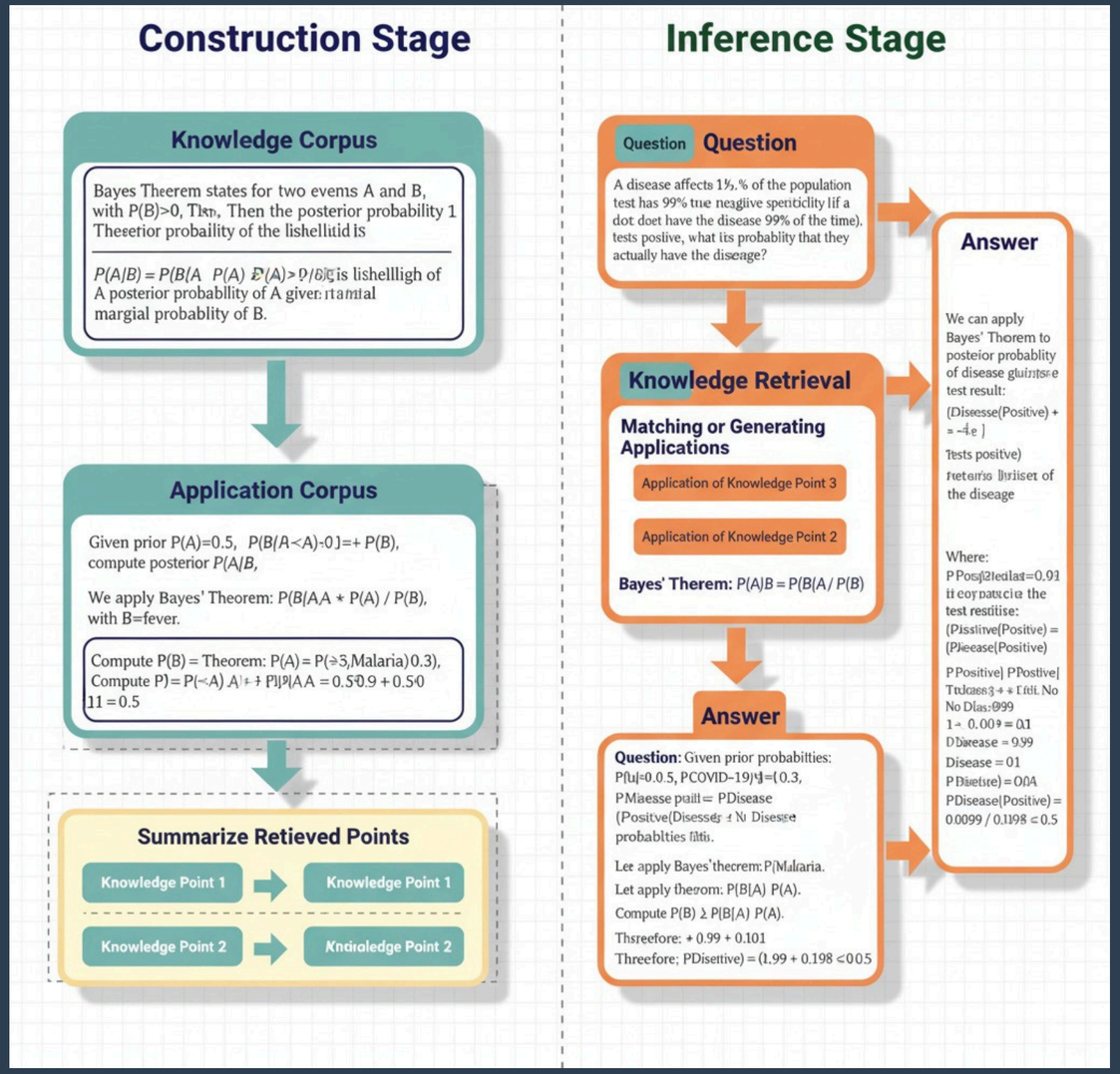
Constructed a domain-specific Legal RAG Embedding Dataset using the axondendriteplus/legal-rag-embedding-dataset framework.
The dataset includes curated case summaries, statutes, and clause mappings, providing high-quality embeddings for improved legal document understanding and retrieval.

Semantic Mismatch Resolution:

Addressed the issue of semantic drift and lexical mismatch by integrating a Top-K semantic retrieval module using embedding-based similarity scoring.
This ensures conceptually similar queries are retrieved even when phrasing or terminology differs, improving the accuracy and robustness of information retrieval.

Architecture Diagram





Benefits of Our Model

Increased Efficiency

RAG+ streamlines data retrieval processes, enabling teams to access vital information quickly. This leads to **significant time savings**, enhancing productivity and allowing staff to focus on core tasks.

Improved Accuracy

Utilizing advanced algorithms, RAG+ minimizes human error in data interpretation. This ensures that users receive **accurate insights**, fostering trust and confidence in decision-making processes across all departments.

Enhanced Decision Making

RAG+ provides *real-time data insights*, enabling organizations to make informed decisions swiftly and accurately, enhancing strategic planning and resource allocation across various sectors.

Transparent and Explainable Results

Every AI-generated legal output is accompanied by a Source Transparency Layer, which lists all the statutes, sections, and case laws referenced, along with relevance scores.
This allows end users to verify how each retrieved source influenced the model's reasoning.

Thank You for Your Attention!

~G448

HARSHITH KULKARNI -23BD1A053T

G. SHARATHCHANDRA REDDY -23BD1A052L

J. ANIRUDH REDDY -23BD1A052T

G. RAGHU VAMSHI gay -23BD1A052N

K. PHANI PRAHAS -23BD1A0533