Madhu Sai Hemanth Nayani

My Portfolio | madhunayani12@gmail.com | (+91) 9392420441 | LinkedIn: madhu-nayani | GitHub: madhunayani

OBJECTIVE

Fuelled by a passion for clean code, AI-driven innovation, and high-performance applications, I aim to accelerate impact as a MERN Stack Developer and AI/ML Enthusiast. With MongoDB, Express, React, and Node as my pit crew — and hands-on expertise in Generative AI, RAG pipelines, and Machine Learning projects — I build full-stack solutions that are not only fast and scalable but also intelligent and future-ready. I'm not here to maintain the status quo — I'm here to engineer the next-level digital experience with velocity, precision, and relentless drive.

SKILLS

Languages & Scripting: Python, JavaScript
Web Technologies: HTML, CSS, Bootstrap, REST API
Frameworks: Node.js, Express.js, React Js
Databases: SQL, RDBMS, Vector Databases (Chroma DB, FIASS DB)
Tools: Git, GitHub, Microsoft Copilot Studio

NLP & Generative AI : BERT, GPT, LLaMA, Lang Chain, LlamaIndex, OpenAI API, Nova Models, Prompt Engineering
Libraries & Frameworks : NumPy, Pandas, Sci-Kit Learn, TensorFlow, Django
Machine Learning & Deep Learning : Scikit-learn, TensorFlow, PyTorch, CNN, RNN, LSTM, GANs, YOLO, Regression, Classification, Decision Tree, Random Forest & Boosting, K-Means Clustering, SVM, Time Series

EDUCATION

Acharya Nagarjuna University — **Guntur, AP**
B.Tech in Computer Science and Engineering — *2020 May - 2024 May*
- **Concentrations:** Computer Science
- **GPA:** 7.0/10.0

Vignana Bharathi Jr. College — **Chirala, AP**
Intermediate — *2018 Mar - 2020 Apr*
- **Concentrations:** MPC(Mathematics, Physics, Chemistry)
- **GPA:** 8.0/10.0

INTERNSHIP

AWS APAC Solutions Architecture virtual experience program on Forage — **December 2024**

- **Designed a scalable web hosting architecture** using **AWS Elastic Beanstalk** to address performance issues caused by increased traffic
- Improved response time and scalability through **auto-scaling**, **load balancing**, and **resource optimization**.
- Delivered a **client-friendly explanation** of the solution, covering how it works and providing an estimate of AWS cost structure.
- Explained complex architecture in a client-facing presentation using visual tools to simplify AWS scalability concepts.

PROJECTS

Legal Document Search using RAG

Description: This project focuses on building a Retrieval-Augmented Generation (RAG) system to enhance semantic

search over structured legal documents. By leveraging Chroma DB as a vector database and sentence-transformers for embedding generation, the system retrieves case laws, contracts, and statutory references based on user queries. The approach improves legal research accuracy and streamlines access to domain-specific information.

Key Contributions:

- Designed a **scalable RAG pipeline** for retrieving legal documents using semantic search.
- Implemented an **embedding-based search mechanism using all-MiniLM-L6-v2** to improve legal context retrieval.
- Integrated **Chroma DB for fast** and **efficient vector storage** and **querying of legal content**.
- Optimized **ingestion of legal metadata** (e.g., jurisdiction, section, clause) for **precise search results**.
- Enhanced relevance ranking of legal documents using cosine similarity and citation context.
- **Integrated Search** and **Blob Storage for scalable indexing** of legal contracts and case archives.

Tech Stack: Python| OpenAI| GPT-3.5 Turbo| Lang Chain| NLP| PDF Processing

Insta Share – Instagram Clone (MERN Stack)

Description: Developed a full-stack social media application replicating Instagram's core functionalities with a scalable MERN stack architecture.

Key Contributions:

- Implemented secure **JWT-based authentication** for user login and session management.
- Built features such as **post creation, likes, comments, and follow/unfollow system** with optimized MongoDB queries.
- Designed a responsive **React frontend** with dynamic feed rendering and state management.
- Integrated **RESTful APIs with Express.js and Node.js** to handle user actions and media interactions.
- Deployed the app ensuring smooth performance and scalability for real-world usage.

Tech Stack: React.js | React Router | Component Lifecycle Methods | Responsive Web Design | REST API Integration | JWT Authentication | Cookies Handling | Authorization & Protected Routes | Figma | Cloudinary | Dynamic Data Rendering | Search Functionality | State Management | CSS

Educational Content Search using RAG

Description:This project focuses on building a Retrieval-Augmented Generation (RAG) system to enhance semantic search over structured educational content. By leveraging Chroma DB as a vector database and sentence-transformers for embedding generation, the system efficiently retrieves relevant content based on user queries. The approach improves response accuracy for downstream models and helps categorize user interests.

Key Contributions:

- Designed a scalable RAG pipeline for retrieving educational content using semantic search.
- Implemented an embedding-based search mechanism using all-MiniLM-L6-v2 to improve retrieval accuracy.
- Integrated Chroma DB for fast and efficient vector storage and querying.
- Optimized data ingestion and metadata management for better organization and retrieval.
- Enhanced search results by using cosine similarity for ranking the most relevant content.
- Integrated Azure Search and Blob Storage for scalable content indexing and retrieval.

Tech Stack: Python | Lang Chain | RAG | Chroma DB | FAISS | Azure AI Search | Azure OpenAI | Azure Cosmos DB