Manasa Kumari

New York | 934-263-2033 | manasa.kumari@stonybrook.edu | linkedin.com/manasakumari | github.com/manasak2010

EDUCATION

Stony Brook University - 3.8/4.0 — Stony Brook, New York
Master of Science in Data Science — August 2024 – Present

Jawaharlal Nehru Technological University Hyderabad - 3.6/4.0 — Hyderabad, India
Bachelor of Technology in Computer Science and Business Systems — June 2020 – July 2024

TECHNICAL SKILLS

Languages: Python, C/C++, Java, SQL, R, MATLAB, JavaScript, HTML, CSS
Data Analysis & Visualization: Tableau, Power BI, Excel, Streamlit, Business Intelligence
Developer Tools: Git, GitLab, GitHub, Agile methodology, REST APIs, CI/CD Pipelines
Libraries: Numpy, Scikit-learn, Keras, Pytorch, Tensorflow, Langchain, Hugging Face, NLTK, Matplotlib, Seaborn
Databases & Services: MongoDB, Stripe, Resend (Email API)
Technologies: MERN Stack, API Development, Machine Learning, Deep Learning, NLP, Computer Vision, Vector Search, Retrieval-Augmented Generation (RAG), Embeddings

EXPERIENCE

Hawl Technologies LLC — Remote, USA
Backend & AI Engineer Intern — May 2025 – Present

- Built and deployed full-stack backend systems using Python and FastAPI, including complete user authentication flow with email verification, password reset, session management, and MongoDB integration.
- Integrated Stripe for secure payment processing and subscription handling.
- Developed data pipelines for web scraping, document chunking, and vector embedding generation using OpenAI and FAISS to support semantic search.
- Applied Retrieval-Augmented Generation (RAG) for intelligent querying over large unstructured datasets.
- Designed a localhost deployment pipeline to enable offline testing and development of AI-driven features.

Blue Cloud Softech Solutions — Hyderabad, India
Software Engineering Intern — February 2024 – April 2024

- Developed a scalable backend system using PHP and SQL to support data warehousing and transformation, integrating with Python pipelines that processed 1,500+ news articles for category tagging and trend analysis.
- Engineered a **multilingual data management system with automated categorization** to dynamically organize article's databse by language and subcategory, leading to a 20% increase in user engagement.
- Designed and deployed an interactive dashboard using Pandas, Matplotlib to analyze user behavior, leading to data-driven improvements in content recommendation and a 20% increase in returning users.

Techsture Technologies India Private Limited — Ahmedabad, India
Machine Learning Intern — August 2023 – October 2023

- Engineered a **vehicle license plate recognition pipeline** using Python, Haar Cascade, and OpenCV, achieving a 40% speed improvement and a 25% reduction in errors.
- Designed and optimized a three-layer CNN model, using TensorFlow to streamline data processing, improving model throughput by 30% and reduced training time; enhancing system efficiency and reducing costs.
- Built a high-accuracy recognition pipeline, achieving a 95% accuracy rate in plate recognition across varied lighting conditions, contributing to the robustness and reliability of the application in real-world scenarios.

PROJECTS

Slipstream | PyTorch, Triton, FlashAttention, CUDA, Mistral-7B, Hugging Face Transformers

- Built custom Triton GPU kernels for fused linear and GELU operations, and integrated Global KV Cache reuse in Mistral-7B, improving token generation efficiency with a 64% increase in throughput and 40% reduction in latency.
- Integrated **speculative decoding** with a fine-tuned TinyLlama draft model, reducing perplexity by over 96% (81.9 → 3.2) across 5,000 QA prompts from the SQuAD v1.1 dataset.

CollegeDataBot | LangChain, BeautifulSoup, Streamlit, Google Cloud Speech-to-Text API, CSV

- Engineered a modular LangChain-based chatbot capable of answering structured queries over scraped college data, with a Streamlit frontend optimized for usability and responsiveness.
- Enabled multimodal interaction by integrating Google Cloud Speech-to-Text for voice input and automating data collection via BeautifulSoup, reducing manual effort by 40%.

Travera | OpenCV, DBSCAN, Firebase, TensorFlow

- Designed a real-time video analysis pipeline using DBSCAN for temporal pattern recognition and unsupervised clustering across video frames, improving retrieval accuracy by 80%.
- Implemented automated data filtering and storage with Firebase, reducing noise and redundant footage by 55% while supporting continuous hands-free data capture.

PUBLICATIONS

- Determining the Optimal Parameters for Wheat and Paddy Crops Using Smart Farming **(IEEE, 2023)** LINK
- Data-Driven Soil Feature Identification in Precision Agriculture for Select Crops **(ICMED-ICMPC 2023)** LINK