

# Manasa Kumari

New York | 934-263-2033 | [manasa.kumari@stonybrook.edu](mailto:manasa.kumari@stonybrook.edu) | [linkedin.com/manasakumari](https://www.linkedin.com/in/manasakumari) | [github.com/manasak2010](https://github.com/manasak2010)

## EDUCATION

### Stony Brook University - 3.8/4.0

Master of Science in Data Science

Stony Brook, New York

August 2024 – May 2026

### Jawaharlal Nehru Technological University Hyderabad - 3.6/4.0

Bachelor of Technology in Computer Science and Business Systems

Hyderabad, India

June 2020 – July 2024

## TECHNICAL SKILLS

**Programming & Data:** Python, R, SQL, C/C++, Java, Go

**Machine Learning & AI:** Scikit-learn, PyTorch, TensorFlow, Keras, Hugging Face, LangChain, Generative AI, LLM Fine-tuning, Prompt Engineering, NLP, Computer Vision

**MLOps & Deployment:** Docker, Kubernetes, MLflow, FastAPI, CI/CD Pipelines, AWS (SageMaker, S3), Firebase

**Data Engineering & Databases:** SQL, MongoDB, ETL Pipelines, Data Lakes, Vector Databases (FAISS)

**Visualization & BI:** Tableau, Power BI, Streamlit, Matplotlib, Seaborn, Excel

**Developer Tools:** Git, GitLab, GitHub, Agile Methodology, REST APIs

**Web & Full-Stack Development:** HTML5, CSS, JavaScript, React, Node.js, API Development, Backend Architecture

## EXPERIENCE

### Hawl Technologies LLC

Remote, USA

Data Science & AI Intern

May 2025 – Present

- Built **LLM-powered data pipelines** with web scraping, embeddings (FAISS, OpenAI), and Retrieval-Augmented Generation (RAG), enabling semantic search across large unstructured datasets.
- Designed and deployed **FastAPI backend using Docker**, implementing secure authentication, CI/CD pipelines and improving offline testing speed by 50% and reducing integration errors.
- Applied MLOps practices (containerization, monitoring, reproducibility) to streamline development and ensure **production readiness**.
- Delivered scalable AI-driven features with supported **real-time user queries** with improved accuracy and lower latency.

### Blue Cloud Softech Solutions

Hyderabad, India

Data Analyst Intern

February 2024 – April 2024

- Designed a **data transformation pipeline** with Python + SQL to process 1,500+ news articles, enabling **automated tagging and trend analysis**.
- Implemented a **multilingual categorization system**, improving content personalization and boosting user engagement by 30%.
- Visualized user behavior with an interactive dashboard using Pandas and Matplotlib to analyze user behavior, leading to 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** with OpenCV and Haar Cascade, achieving 95% accuracy under varied conditions.
- Optimized a **CNN model with TensorFlow**, improving throughput by 30% and reducing training time by 25%.
- Enhanced pipeline robustness by improving recognition speed by 40%, ensuring reliability in real-world deployments.

## PROJECTS

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

- Engineered a **high-performance LLM inference pipeline** with Triton GPU kernels, Global KV Cache reuse, and FlashAttention, boosting throughput by 64% and cutting latency by 40%.
- Integrated **speculative decoding** with a TinyLlama draft model, reducing perplexity from 81.9 → 3.2 across 5,000+ QA prompts, enabling **production-scale LLM serving**.

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

- Built a **LangChain-based chatbot** with Streamlit UI to handle structured queries over scraped college datasets, reducing manual lookup time by 40%.
- Integrated **multimodal interaction** via Google Cloud Speech-to-Text, enabling voice-driven queries and improving accessibility for end users.

### Travera – OpenCV, DBSCAN, Firebase, TensorFlow

- Developed a **real-time video analysis pipeline** using DBSCAN clustering for temporal pattern recognition, improving retrieval accuracy by 80%.
- Automated **data filtering and storage with Firebase**, cutting redundant footage by 55% and enabling 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](#)