

# SAKSHAM PATTEM

Tempe, Arizona

☎ (934) 221-1909 ✉ [saksham.pattem@gmail.com](mailto:saksham.pattem@gmail.com) [in](#) [Linkedin](#) [G](#) [Github](#)

## Education

**Stony Brook University | GPA: 3.50/4.0**

**Aug 2022 – May 2024**

Stony Brook, NY, USA

- Master of Science in Computer Science
- Coursework: Machine Learning, Fundamentals of Computer Networks, Theory of Databases

## Technical Skills

**Programming Languages**: C++, Java, Javascript, Python, Golang, SQL, MongoDB

**Frameworks and Libraries**: Django, Flask, FastAPI, React, Node.js, Express, REST

**Technologies and Tools**: Git, Docker, Kubernetes, AWS, GCP, Hadoop, Spark, PySpark, LlamaIndex

## Professional Experience

**Redclay Consulting**

**Jul 2024 – Present**

Software Consultant

Jersey City, NJ, USA

- Integrated Oracle Work and Asset Cloud Service (WACS) with Oracle Field Service (OFS) using Oracle Integration Cloud, enabling real-time synchronization of asset and field data for a major utility provider.
- Engineered client-specific plugins using reusable, framework-agnostic JavaScript components for OFS, and developed tailored business logic extensions for WACS using Groovy.
- Built an internal Intelligent Meeting Insights Engine using Retrieval-Augmented Generation (RAG), streamlining access to project requirements from meeting transcripts and reducing Mean Time to Retrieve (MTTR) by 30%.

**Biomedical Informatics, Stony Brook University**

**Jun 2023 – Apr 2024**

Research Assistant

Stony Brook, NY, USA

- Integrated WSInfer with QuPath by automating the creation of visualization-ready projects from GeoJSON predictions, enabling overlay of detection results for expert review.
- Implemented patch-wise test-time augmentation to improve model robustness on diverse image orientations, increasing tumor and lymphocyte detection accuracy by 8% on whole slide images.
- Dockerized the CLI tool for cross-platform deployment and developed extensive PyTest suites with 95% edge case coverage, ensuring reliable performance across model backends and speed configurations.

**École De Technologie Supérieure**

**Jul 2021 – Sep 2021**

Mitacs Summer Research Intern

Montréal, QC, Canada

- Optimized image preprocessing with an AWS S3 + Lambda automation pipeline, triggering validation and transformation on new uploads, reducing manual errors and preprocessing time by 50%.
- Boosted inference throughput by decoupling GPU workloads with async task queues and batching, then parallelizing execution across Docker containers orchestrated via Kubernetes—achieving a 45% speedup on large image sets.
- Delivered real-time inference results and patient metadata via WebSocket-based communication, enabling integration with visualization tools and improving interactivity for radiologists.

**Verzeo**

**May 2020 – Jun 2020**

AI Software Engineer Intern

Chennai, TN, India

- Built and deployed an edge-computing solution using OpenCV and YOLOv3 to monitor office occupancy, reducing close-contact risks during the pandemic.
- Developed a Python-based REST API using Flask to serve real-time people counting data from Raspberry Pi devices; enabled remote monitoring via lightweight HTTP endpoints.
- Packaged the system into a deployable unit using systemd + shell scripts for auto-start and recovery, and optimized image processing throughput on limited hardware (Raspberry Pi 3B+) while maintaining 88% detection accuracy.

## Projects

**Custom DNS Resolver** | *Python, DNSPython*

- Built a DNS resolver in Python to understand DNS resolution flow and how domain names are translated to IP addresses.
- Implemented features like iterative and recursive lookups, with logging for educational and debugging purposes.

**Geostatistical Crime Prediction using PySpark** | *PySpark, Hadoop, Machine Learning*

- Built a scalable crime prediction pipeline using PySpark on Hadoop, processing over 1M records for real-time analysis.
- Conducted a comparative study of K-Means, Naïve Bayes, and Linear Regression models; selected the optimal model based on accuracy, runtime, and false positive rate—reducing false alarms by 25% and improving precision by 60%.

**PodPilot - Podcast Curation Agent** | *LangChain, ChromaDB, Claude, React*

- Developed an AI-driven podcast insight engine using LangChain and ChromaDB to surface relevant content and enable efficient cross-podcast knowledge curation, improving content discovery for users.
- Leveraged Claude for response generation, conditioning on retrieved podcast context and enhancing user engagement with related-question suggestions, reducing information retrieval time by 3 hours per podcast.

## Fellowships

**Globalink Graduate Fellowship** : Received \$15000(CAD) support for pursuing graduate studies from Canada, offered to only about 300 undergraduate students from 16 countries.