

DIYA SAHA

☎ (408)-480-7445 ✉ diya.cottonian@gmail.com [in diya-saha](#) [DiyadotSaha](#)

Education

University of California, Irvine

Sep. 2023 – Jun. 2025

Master of Science in Computer Science

University of California, Santa Cruz

Sep. 2019 – Jun. 2023

Bachelor of Science in Computer Engineering with specialization in Robotics; Minor in CS

Experience

Canine Cancer Research with Dr.Pierre Baldi

July 2024 – Present

Graduate Researcher

Irvine, CA

- Analyzed various clinical reports of canine subjects in the Vidium Animal Health database, using machine learning models to study patterns in detected biomarkers
- Utilized LLMs to generate automated summaries and overviews of the clinical reports, aiding in the automation of the clinical report generation process

DSights.Inc

Summer 2022

Software Research Intern

Remote

- Used Natural Language Processing Techniques to automatically compare restaurant menus from web pages
- Explored different ideas for a prototype to help our client find similarities between restaurant menus based on words and phrases and submitted a written report summarizing the findings
- Conducted analysis on the menus publicly available and explored different methods to visualize data using Python libraries like Matplotlib, Pandas and Seaborn, and Google sheets

Tech4Goods Labs

April 2021 – September 2022

Undergraduate Research Assistant

Santa Cruz, CA

- Analyzed different methods of youth career mentorship through the organization, Your Future Our Business(YFIOB). Built low and high-fidelity prototypes of portals using Figma
- Worked efficiently in a team setting to create a grant study proposal to be presented to the Institutional Review Boards(IRB) to receive approval to conduct the aforementioned research

Projects

Context-Aware Chatbot using RAG | Langchain, FAISS, HuggingFace

June 2024

- Developed a context-aware chatbot tailored for UCI students, facilitating course inquiries and providing context-specific responses
- Executed end-to-end machine learning pipeline from creating a vector database to create a prompt template to help the chatbot answer questions
- Link to GitHub repository

Determining readmission rate on diabetic patients | Python, Pandas, Matplotlib, Sklearn

September 2023

- Utilized a comprehensive decade-long dataset from 130 US hospitals to analyze patterns related to the early readmission of diabetic patients, employing classification, ensemble, and neural network models.
- Implemented data imputation, oversampling techniques to address class imbalance, and applied feature selection for model optimization.
- Link to GitHub repository

UAV Model Simulator, Evade and Avoid | Python, GUI, PyQt5, Matplotlib

March 2023

- Developed and integrated evasive maneuvers functionality into the flight simulator, implementing new control laws to ensure avoidance of explosive forces along the aircraft's trajectory.
- Enhanced autopilot model by incorporating a robust state machine and refined waypoint following algorithm, resulting in visually and computationally effective aircraft maneuvers while avoiding potential explosive hazards.
- Link to GitHub repository

Technical Skills

Proficient in Programming Languages: Python, Java, C/C++, HTML/CSS, JavaScript, SQL, Assembly

Developer Tools: VS Code, Eclipse, Google Cloud Platform, Amazon Developer Tools, Vivado, Figma, GitHub

Technologies/Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, Langchain, HuggingFace

Honors / Certificates

- Received Dean's Honor on multiple quarters for being in the top 15% of my class.
- Google Professional Certification for Data Analytics, Google Professional Certification for UI/UX Design
- Completed Coursera Machine Learning Specialization by Andrew Ng and Google's Introduction to Generative AI