**Education**

**Masters of Computer Science Sept 2022 – June 2024**

University of California San Diego (UCSD) CGPA: 3.979/ 4

*Relevant Courses*: Deep Learning, Scalable ML Systems, Recommender Systems, Computer Vision, NLP – LLMs, Algorithms

**Bachelor Of Engineering (Computer Engineering) August 2018 – July 2022**

Vivekanand Education Society’s Institute of Technology (VESIT) CGPA: 9.013/ 10

*Relevant Courses*: Machine Learning, Software Development, Natural Language Processing, Cloud Compute, Big Data Analytics

**Experience**

*Software Engineer (Data and AI),* **Ujima S&P Lab, UCSD**  **March 2023 – Present**

* Building auto scaling **AWS Cloud architecture** to host an LLM User Study on **AWS Beanstalk** with a **Mongo DB** storage system.
* Led the **product development** and strategy for an LLM orchestration platform, managing a team of 3 to curate a structured, analysis-ready dataset for HCI research on hate crimes; utilized **LLMs**, **clustering**, and **NER** to **web-scrape**, merge, and process articles from sources like the Web, Lexis Nexis, and TDM Studio, employing LLMs for filtering, **classification**, and **data extraction**.
* Built a “Smart Mirror” on **Raspberry Pi** using the **VGG-Face** TF-Lite Model, trained on 100,000 images, to detect ethnicity, highlight the existing biases in CV datasets, and contribute to the **public domain** via user feedback.

*Cloud Cyber Security & Data Analyst Intern,* **Legendary Entertainment**  **June 2023 – Dec 2023**

* Create a **Python script** automating data flow to generate dashboards, **reducing time taken by the SOC analyst from 80 minutes** to about **4 minutes daily** in classifying user-anomaly events by integrating **Splunk**, **Active Directory**, and **Azure MSGraphs APIs**.
* Part of **cross-functional** collaborative team to maintain a high level of security while streamlining users’ **authentication** process by integrating a **FIDO Alliance** product into the existing **SAML/OIDC SSO** workflow.
* Assisted the **executive leadership** in laying the foundation of a next-gen universal **anomaly** and user-behavior detection platform based on integrating LLMs with existing cloud brokers in collaboration with **Sky High Security** by building a prototype in **Splunk**.

*Full Stack Development Intern,* **Makos Infotech (Startup)**  **June 2021 – August 2021**

* Managed the **end-to-end development** of an internal platform **automating** on-campus placement processes for universities.
* Applied **Agile** methodologies, coordinating with stakeholders to deliver features that improved **UX** and **operational efficiency**.
* Maintained and optimized **relational database** storage using **MySQL** workbench and deployed it on **AWS RDS** to develop a college-student-company social network inspired by Facebook’s friend system.

*Undergraduate Research Assistant,* **Tata Institute of Fundamental Research (TIFR)**  **June 2021 – May 2022**

* Led a team of 4 to develop an **Android Java** application that monitors a selected directory and uses **multi-part upload** methodologies to **encrypt** and **securely** upload to a dedicated private remote cloud server.
* Published a [paper](https://doi.org/10.1007/978-3-031-18497-0_41) explaining our Node JS based **Fault Tolerant** **client server architecture** connected to remote stations.
* Utilized **GCP’s Maps** and **Sheets API** to design a real time **HTML/CSS** based live **Geo tracking** website from the collected data.

*Data Analyst Intern*, **Leadingindia.ai** **May 2020 – July 2020**

* Collaborated with a team of four to develop a **vaccine prediction model** for H1N1 and seasonal flu vaccines, accurately predicting public acceptance trends (41%) for the COVID-19 vaccine, securing **first place** among 85 intercollege peer groups.
* Published a [research paper](https://doi.org/10.1007/978-981-16-0401-0_11) in Springer & authored a [blog](https://medium.com/@jjhaveri1906/pandemics-a-harsh-reality-7c05254e907b) highlighting the **correlation** between H1N1 and COVID-19 pandemics.

**Projects**

**[Alt Bot for Mastadon: An automatic image alt generation bot](https://github.com/CSE210-Fall23-Team2/AltBot)** [(Presentation)](https://docs.google.com/presentation/d/1smZzOd8u-NhgbotJRkn2Eqw5WmXVxA-XSB0afzxaNWE/edit?usp=sharing) **Sep 2023 – Dec 2023**

* Developed a **REST-API** based **Chrome extension** in **JS** to help visually impaired people browse decentralized social media feeds by leveraging **hugging face** **image captioning** models to generate alternative image descriptions and **injecting** them in **HTML**.
* Deployed 3 levels of **custom cache system** to ensure **efficient** performance with **minimal lag** complemented with **testing** scripts.

[**MedLM: Exploring Language Models for Medical QnA Systems**](https://github.com/JayJhaveri1906/CSE291_MedLM)[(Paper)](https://arxiv.org/abs/2401.11389) **March 2023 – Aug 2023**

* Led a team of 4 to **fine tuning** language models (**Bloom, T5, GPT-2**) on the MedQuad Healthcare Doctor-Paitient QnA dataset in collaboration with **Microsoft researcher** Dr. Asma Ben Abacha.
* Compared performance against **GPT-3.5** and **GPT-4** using **Dynamic Prompting** with Retrieval Augmented Generation (RAG) via medical **InstructOR Embeddings** on the patient questions.
* Increased the **ROUGE** and **BLEU** scores by 10% using a **Bert Classifier** to give extra **contextual awareness** to the models.

**[Divya-Drishti: An Independent Aid for the Visually Impaired](https://github.com/JayJhaveri1906/Divya-Drishti)** [(Paper)](https://dx.doi.org/10.2139/ssrn.3867707) **Aug 2020 – May 2021**

* Achieved a **400% net cost reduction** by creating a real-time **Voice-activated AI IoT** android application to help Visually Impaired People (VIPs) comparable to state-of-the-art OrCam in **multimodal classification** of currency, objects utilizing **GCP’s Vertex AI**.
* Published a [research paper](https://dx.doi.org/10.2139/ssrn.3867707) highlighting the needs of VIPs, funded by the **Mumbai University Minor Research Grant**.

[**Automated Number Plate Recognition and Parking System**](https://github.com/JayJhaveri1906/AutomaticParkingSystemANPR) **Dec 2019 – Feb 2020**

* Used **CI/CD** to build an **Android Java** based application connected to a **Firebase server** to automate security and space availability in car parking systems by monitoring the number plates **detected** at the entry and exit using **Tesseract OCR** and **YoloV3**.
* **Containerized** the software to ensure **seamless deployment** on existing CCTVs systems at parking lots, **minimizing** upfront **costs**.

**Selected Certifications and Publications**

- *AWS Certified Cloud Practitioner (CCP)*, Amazon Web Services, [Credly Link](https://www.credly.com/badges/45871348-ade7-4b07-89ee-6eeb7e85b72f/linked_in_profile)

- *Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination*. Advanced Computing. IACC 2020. Springer. [DOI.org Link](https://doi.org/10.1007/978-981-16-0401-0_11)

**Skills:** Python, SQL, PyTorch, Tensorflow, Pandas, HTML, CSS, Javascript, Java, C, Docker, Splunk, Linux, Git, AWS, GCP, Firebase