JAY JHAVERI

(858)214-9192 | jjhaveri@ucsd.edu | San Diego | https://jayjhaveri190600.web.app/

EDUCATION

Masters of Computer Science   **Sept 2022 – June 2024**
University of California San Diego (UCSD)   CGPA: 3.972/ 4
  Relevant Courses: Deep Learning, Scalable Data/ML Systems, Recommender Systems, Computer Vision, Adv NLP - LLMs

Bachelor Of Engineering (Computer Engineering)   **August 2018 – July 2022**
Vivekanand Education Society's Institute of Technology (VESIT)   CGPA: 9.013/ 10
  Relevant Courses: Software Development, DBMS, Algorithms, Machine Learning, Object Oriented Programming, Big Data

EXPERIENCE

Machine Learning NLP and CV, **Ujima S&P Lab, UCSD**   **March 2023 – Present**
- Led 3 freshers through the Early Research Program in curating a database using LLMs, Web scraping, and NER for HCI Research.
- Built a "Smart Mirror" on Raspberry Pi using REST API and 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.

Data Analyst & Cyber Security Intern, **Legendary Entertainment**   **June 2023 – Dec 2023**
- Created a Python automation script to generate a dashboard that reduced the total time taken by the SOC analyst from 80 minutes to about 4 minutes daily in tracking user-anomaly events by integrating Splunk Dashboard and Azure MSGraphs APIs.
- Maintained a high level of security while streamlining a user's authentication process by integrating a FIDO Alliance product into the existing SAML/OIDC SSO workflow.
- Assisted the VP of security in laying the foundation of a next-gen universal anomaly and user-behavior detection platform based on LLMs in collaboration with Sky High Security by building a prototype in Splunk.

Full Stack Development Intern, **Makos Infotech (Startup)**   **June 2021 – August 2021**
- Integrated and developed server-side CRUD code using JQuery, PHP, and MySQL for an early startup targeting the automation of the On-campus placement process using Agile Methodologies.
- Created and merged relational databases using MySQL workbench and deployed it on AWS RDS to develop a college-student-company social network inspired by Facebook's friend system.
- Established a mentorship-onboarding program for new undergraduate interns, aligning them with the existing codebase and processes, saving the company at least 1 week of time and effort.

Web Developer Intern, **VESIT Renaissance Cell**   **June 2020 – July 2020**
- Led the design and development of a Django-based Paper Publication Web service, implementing essential CRUD operations for 50 professors hosted on Heroku, serving approximately 250 CSE students.
- Developed a Portfolio Website template using HTML/CSS/JS and Bootstrap, deployed on Google's Firebase.

PROJECTS

Inquirable Models: Increasing Explainability in Health-AI using LLM   **Sep 2023 – Jan 2024**
- ✓ Conducted a two-phase exploratory study using prompt engineering techniques on leading Large Language Models (LLMs) with SHAP values to improve the interpretability of traditional medical risk models and reduce patient risk.
- ✓ Facilitated surveys with doctors and patients to assess answer quality, focusing on metrics such as confabulation rate.
- ✓ Paper's poster was accepted for presentation at the AMIA 2024 Annual Symposium.

Alt Bot for Mastadon: An automatic image alt generation bot (Github)   **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.
- ✓ 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 (Github)   **March 2023 – Aug 2023**
- ✓ Led a team of 4 to fine tuning language models (Bloom, T5, GPT-2) on the MedQuad 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 Documentation (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.

Aatmanirbhar Sanchar: Secure Self-Sufficient Communications (Github)   **June 2021 – May 2022**
- ✓ Led a team of 4 to develop a secure, end-to-end encrypted chat application with off-the-grid functionality for the Tata Institute of Fundamental Research Institute (TIFR), utilizing CI/CD practices to establish a streamlined client-server architecture.
- ✓ Collaborated in implementing SHA-256 and AES-256 overlapped inside an HMAC envelope to protect against cyber attacks.

Divya-Drishti: An Independent Aid for the Visually Impaired   **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 highlighting the needs of VIPs, funded by the Mumbai University Minor Research Grant.

SELECTED CERTIFICATIONS AND PUBLICATIONS

- *AWS Certified Cloud Practitioner (CCP)*, Amazon Web Services, Credly Link
- *Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination*. Advanced Computing. IACC 2020. Springer. DOI.org Link

Technical Skills: Python, HTML, CSS, JavaScript, SQL, PHP, React, REST, PyTorch, Selenium, Splunk, C, Java, Git, AWS, Google Cloud