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: 3.97/ 4
Relevant Courses: Deep Learning, Bioinformatics / Genetics, Scalable Data/ML Systems, Computer Vision, Advance 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: Machine Learning, Software Engineering, Database Management, Data Warehouse, Data Analysis, Java

INTERNSHIP EXPERIENCE

Machine Learning NLP, CV Researcher, **Ujima S&P Lab, UC San Diego** **March 2023 – Present**
- Building a "Smart Mirror" on Raspberry Pi using VGG-Face TF-Lite Model, trained on 100,000 images, to detect ethnicity, address the existing biases in CV datasets, and contribute to the public domain via user feedback.
- Led 4 undergraduates through the Early Research Student Program in analyzing privacy data for MMO AR/VR games using TF-IDF and K-means Clustering.

Data Analyst & Cyber Security Intern, **Legendary Entertainment** **June 2023 – Dec 2023**
- Created a shortcut 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 via Python Scripting.
- 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 code using JQuery, PHP, and MySQL for an early startup targeting the automation of the On-campus placement process 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 interns, saving the company at least 1 week of time and effort.

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 in Springer & authored a blog highlighting the correlation between H1N1 and COVID-19 pandemics.

PROJECTS

Global Ancestry Comparision (Bioinformatics) (Github) **Jan 2024 – March 2024**
- ✓ Led the development and benchmarking of multiple dimensionality reduction techniques such as PCA, t-SNE, UMAP, MDS, Isomap, and Autoencoders for global ancestry inference on 1000 Genome project's VCF files.
- ✓ Awarded the class best presentation for indepth analysis by using homogeneity scores to compare the techiques.

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 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-tune 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.

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 product OrCam in using machine learning to detect currency, objects, and scenes.
- ✓ Published a research paper highlighting the needs of VIPs, funded by the Mumbai University Minor Research Grant.

SELECTED RESEARCH PUBLICATIONS

Inampudi S., **Jhaveri J.** et al., (2021) **Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination**. Advanced Computing. IACC 2020. Communications in CIS, vol 1367. Springer, Singapore. DOI.org Link

Skills: Python, PyTorch, Tensorflow, OpenCV2, Hugging Face, Java, Pandas, Numpy, SciPy, SQL, C, Linux, Git, AWS, Azure, Firebase