JAY JHAVERI

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

EDUCATION

Masters of Computer Science **Sept 2022 – June 2024 (Expected)**
University of California, San Diego (UCSD) CGPA: 3.97/ 4
Relevant Courses: Advanced NLP via LLMs, Deep Generative Modelling, DL, Scalable ML Systems, Computer Vision

Bachelor Of Engineering (Computer Engineering) **August 2018 – July 2022**
Vivekanand Education Society's Institute of Technology (VESIT) CGPA: 9.013/ 10
Relevant Courses: Natural Language Processing, Machine Learning & AI, Big Data Analytics, Software Development

INTERNSHIP EXPERIENCE

Cyber Security & Machine Learning Intern, **Legendary Entertainment** **June 2023 – Dec 2023**
- Reduced false positive anomaly detection time by 10x by streamlining employee online activity monitoring using Splunk Dashboard, Python scripting and integration with Azure MSGraph API.
- Contributed to integrating a FIDO Alliance product into the SSO workflow, enhancing security and user experience.
- Assisted in foundational work for the Shared Learning Intelligence Platform (SLIP) to improve anomaly detection in security cloud brokers in collaboration with Sky High Security.

Full Stack Development Intern, **Makos Infotech** **June 2021 – July 2021**
- Developed Server-side rendering for their main website (Jobaskit.com) utilizing JQuery, PHP, and MySQL, which targets automating the On-campus placement process for various colleges.
- Managed existing and created relational databases using MySQL Workbench and deployed them on AWS.
- Co-pitched the online job placement portal, Jobaskit, to 3 University professors alongside the founder.
- Mentored 2 intern recruits working on the digitalization of the teaching process.

Data Analyst Intern, **Leadingindia.ai** **May 2020 – June 2020**
- Worked in a team of four to build a Vaccine Prediction model on the H1N1 and seasonal flu vaccines to accurately predict the trends of the public acceptance rate (41%) of the Covid-19 vaccine.
- Research Paper was published in Springer & I wrote a Blog showcasing the correlation between the two pandemics.
- Secured first position for the mentioned research project amongst 85 peers intercollege.

PROJECTS

Inquirable Models: Increasing Explainability in ML using LLM **Sep 2023 – Jan 2024**
- ✓ Explored the possibility of making traditional medical risk models more easily interpretable using Large Language models with the help of SHAP values, ultimately reducing the patient's risk.
- ✓ Conducted exploratory research with the help of prompt engineering on popular LLMs in a 2 stage manner.
- ✓ Hosted surveys for Doctors and Patients to evaluate the answers generated on metrics such as Confabulation rate.

MedLM: Exploring Language Models for Medical QnA Systems **March 2023 – Aug 2023**
- ✓ Led team of 4 in fine-tuning diverse language models (e.g., bloom, t5, gpt2) on the MedQuad dataset, comparing them with larger models (gpt3.5, gpt4) using direct questions and dynamic prompt engineering.
- ✓ Collaborated with Microsoft researcher Dr. Asma Ben Abacha, creator of MedQuad dataset, for expert guidance.
- ✓ Utilized ROUGE, BLEU metrics and conducted human surveys for doctors and patients to evaluate the model.

GrooveGenie: A copyright-free music generator **March 2023 – June 2023**
- ✓ Created an open source music generation model, utilizing Facebook's EnCodec Transformer model to compress audio wav files to an embedding that can be understood by the model.
- ✓ Training a conditioned GAN network that generates music based on user-provided genre inputs embedded using the BERT model, with a goal of creating only copyright and royalty-free music, being trained on the FMA dataset.
- ✓ Trying out different, more efficient Diffusion/Transformer architecture to generate audio.

Code for Change Hackathon: A Data Extraction project **Nov 2020 – 24 Hours**
- ✓ Developed Django based data extracting software for Global Parli Foundation NGO to automate the translation of Land ownership papers' pdf originally in Devanagari Script into an editable Excel sheet using Google Cloud OCR.
- ✓ Secured First position for the web application amongst the 72 teams participating.

SELECTED RESEARCH PUBLICATIONS

Jhaveri, J., Gupta, A., Chhabria, P., Ochani, N. and Sengupta, S., 2021. **Divya-Drishti: An Independent Aid for the Visually Impaired**. *SSRN Electronic Journal*. DOI.org Link

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

- **Technical Skills:** Python, PyTorch, TensorFlow, OpenCV, MS Office, Splunk, Git, Azure, AWS, Google Cloud, Firebase