**Education**

**Masters of Computer Science Sept 2022 – June 2024 (Expected)**

University of California, San Diego (UCSD) CGPA: 3.97/ 4

*Relevant Courses*: Deep Learning, Scalable 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*: Machine Learning, Software Development, Big Data (Hadoop), Algorithms, Cloud Computing

**Internship Experience**

*Machine Learning & Cyber Security 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 the lead SOC analyst 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.

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

* Worked with 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](https://doi.org/10.1007/978-981-16-0401-0_11) was published in Springer & I wrote a [Blog](https://medium.com/@jjhaveri1906/pandemics-a-harsh-reality-7c05254e907b) showcasing the correlation between the two pandemics.
* Secured first position for the mentioned research project amongst 85 peers intercollege.

**Projects**

**[Alt Bot for Mastadon: An automatic image alt generation bot](https://docs.google.com/presentation/d/1smZzOd8u-NhgbotJRkn2Eqw5WmXVxA-XSB0afzxaNWE/edit?usp=sharing) Sep 2023 – Jan 2024**

* Developed a Chrome extension to help the visually impaired browse decentralized social media by leveraging hugging face image captioning ML models to generate alternative descriptions for images in posts in a CI/CD format.
* Deployed 3 levels of custom cache system to ensure peak and efficient performance with no lag.
* Being part of a team of 10 members, we also envisioned this being useful in easing the search for specific media.

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

* Led a 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](https://github.com/JayJhaveri1906/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.
* Trained 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.

[**Game Genre and Recommendation Classification using Steam Reviews**](https://github.com/JayJhaveri1906/Game-Genre-and-Recommendation-Prediction) **Nov 2022 – Dec 2022**

* Designed data pipelines to preprocess and apply machine learning techniques to classify the game’s genre, user’s sentiment and finally curated a personalized game recommendation system using user reviews.
* Achieved 90.53% accuracy with RF, balanced data & TF-IDF, outperforming N-Gram, Multinomial NB, Linear SVC.

[**Divya-Drishti: An Independent Aid for the Visually Impaired**](https://github.com/JayJhaveri1906/Divya-Drishti) **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 detecting currency, objects, and scenes.
* 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.

**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](https://doi.org/10.1007/978-981-16-0401-0_11)

**Technical Skills:** Python, PyTorch, TensorFlow, OpenCV, Hugging Face, Splunk, Sql, Spark, Docker, Git, Azure, AWS, GCP