**Education**

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

University of California, San Diego (UCSD) CGPA: 3.95/ 4

*Relevant Courses*: Recommender Systems, Adv Computer Vision, Advanced NLP, DL, Scalable Data/ML Systems

**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, Big Data Analytics, AI, Software Development, OOPs (Java), Data Sructure (C)

**Internship Experience**

*Machine Learning & Cyber Security Intern,* **Legendary Entertainment**  **June 2023 – Present**

* Reduced false positive anomaly detection time by 10x by Streamlining employee online activity monitoring using Splunk Dashboard and Python scripting.
* 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](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**

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

* Experimented by comparing Fine-tuned distilled generative text models like GPT2, Bloom with larger general models like GPT 3.5 and 4 for a Doctor Patient QnA conversation.
* 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 prompt engineering.
* Collaborated with Microsoft researcher Dr. Asma Ben Abacha, creator of MedQuad dataset, for expert guidance.
* Utilized ROUGE, BLEU metrics, and conducted user surveys for doctors and patients to evaluate model performance.

**[GrooveGenie: A copyright-free music generator](https://github.com/JayJhaveri1906/GrooveGenie-A-copyright-free-music-generator) March 2023 – Present**

* 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.

[**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 game genres and also personalize game recommendations using the user’s reviews and hours played.
* Out of N-gram, Multinomial NB, and Linear SVC, RF with Balanced data & TF-IDF gave the highest accuracy of 90.53%.

[**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 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.
* Received feedback, on the android-Java app developed, by National Association for the Blind (NAB)’s members.
* Published a [research paper](https://dx.doi.org/10.2139/ssrn.3867707) highlighting the needs of VIPs.

**Research Publications**

Inampudi S., **Jhaveri J.** et al., (2021) **Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination**. In: Garg D., Wong K., Sarangapani J., Gupta S.K. (eds) Advanced Computing. IACC 2020. Communications in Computer and Information Science, vol 1367. Springer, Singapore. (<https://doi.org/10.1007/978-981-16-0401-0_11>)

* **Technical Skills:** Python, PyTorch, TensorFlow, OpenCV2, Java, C, Sklearn, SQL, SPL, Git, Project Management, AWS