Pranav Agarwal

Irvine, CA | (949) 844-0408 | Portfolio | pranav.agarw@gmail.com | Github | Linkedin

EDUCATION:

University of California, Irvine (UCI) | Irvine, CA — Sep 2023 - Dec 2024
Master of Data Science | GPA: 3.9

- Machine Learning; Artificial Intelligence; Bayesian Inference; Statistics; Deep learning

Vellore Institute of Technology, Vellore | Vellore, India — Jul 2017 - Jun 2021
Bachelor of Technology, Computer Science and Engineering | GPA: 9.06

- Data Structures; Database Management; Natural Language Processing; Programming.

SKILLS, CERTIFICATIONS and ACHIEVEMENTS:

- Python; C++; R; SQL; AWS; Docker; Kubernetes; PyTorch; Tensorflow; Spacy; Tableau; LLM; HuggingFace; Langchain; OpenCV; Seaborn; MilvusDB; Neo4J; Seurat; Pandas; Keras; Numpy
- AWS - solutions architect associate
- Applied Machine Learning in Python, University of Michigan – Coursera.
- Open-source contributor of Mozilla - github.
- Student ambassador of the Master of Data Science program - represent UCI ICS and MDS in conferences, talks and events.

WORK EXPERIENCE:

Machine Learning Infrastructure intern | Safran — Jul 2024 - Present

- Collaborated in the design and development of cloud-based data applications, specializing in **machine learning models** for **predictive health management systems** in **aerospace equipment**. Contributed to the transition from proofs-of-concept to functional prototypes.
- Engineered and integrated software solutions to enhance **prototype systems**, verifying their **functional performance** and optimizing **predictive algorithms** to **forecast equipment health** within an **IoT** framework on AWS using **Sagemaker** and **Amazon ECR**.
- Worked within a dynamic, **cross-functional** team to seamlessly merge various technologies, ensuring the robust deployment of **cloud-hosted machine learning solutions** tailored for **real-time health monitoring** in the **aerospace sector**.

Machine Learning Student Researcher | University of California, Irvine — Jun 2024 - Present

- Executed a comprehensive evaluation in the **AI Center of UCI** in **Dr. Jana Lipkova's OctoPath** lab of patch-level and **whole-slide models** (ResNet50, UNI, CONCH, REMEDIS, HIPT) for survival prediction using datasets from **TCGA, CPTAC**, and private cohorts.
- Analyzed model performance, generalization to external data, and interpretability - targeting **brain and lung cancers**.
- Delivered insights into the applicability and effectiveness of foundational models in **clinical survival prediction**, enhancing the understanding of their utility **beyond classification tasks**.

NLP Student Researcher | University of California, Irvine — May 2024 - Present

- Engineered a program in the **UCI INCHES Lab** with **Dr. Angela Lukowski's** using Python and **spaCy** to automatically **parse narrative data** into propositional phrases, enhancing the efficiency of analyzing event memory studies.
- Developed and implemented flexible rules to accommodate varied subjects and verbs within the narratives, ensuring the program's adaptability to different datasets. Reduced manual work of **2 weeks to 1 day**.

Cloud Engineer | Airbus — Jul 2021 - Aug 2023

- Collaborated within a 15-member team to deliver secure and scalable cloud services on **AWS**, facilitating the migration of 65% of departments from on-premises to cloud infrastructure.
- Employed **analytical Bayesian methodologies** to detect and address potential security threats, reducing monthly security alerts by 20% and ensuring **data integrity**.
- Integrated a **recommendation engine** into production, utilizing historical usage patterns to optimize menu offerings, resulting in annual cost savings of **$70,000**.
- Engineered a comprehensive **dashboard** via **Amazon QuickSight**, synthesizing user data metrics to furnish actionable insights for informed decision-making, project management and strategic planning.

NLP Software Engineer Intern | Novartis — Jan 2021 - Jun 2021

- Engineered a **chatbot** via **natural language processing** techniques, reducing customer service response time by 1.7 hours, replacing L1 level support with chatbot saving **$1 million** via bot communication.
- Innovated a **chatbot analyzer** capable of identifying areas of low performance and providing actionable insights for enhancement, contributing to a 30% surge in chatbot usage.
- Performed **A/B testing** on intents via user feedback and incorporated them into language model. Increased positive feedback by 25%.

PROJECTS:

Forest Fire Detection Using Classifiers and Transfer Learning | github | IEEE

- Engineered a **Machine Learning model** for forest fire detection, mitigating the inefficiencies of traditional hardware devices.
- Applied **transfer learning** from pre-trained models like **Inception** and **VGG19** to extract image features.
- Incorporated ML classifiers such as **SVM**s, **KNN, Naïve Bayes** for the prediction, enhancing detection accuracy and response times.

Duplicate Detection in Job Postings using NLP and Milvus | github

- Implemented **Sentence Transformers** to generate **embeddings** from job descriptions, enhancing duplicate detection capabilities by capturing **semantic similarity** and context-aware representations.
- Configured and deployed a Milvus instance, executed vector indexing, and developed a **cosine similarity-based search method** to identify duplicate job postings. Assessed effectiveness using **precision, recall, and F1** score metrics, optimizing similarity thresholds.

Bombay Stock Exchange Equity Analysis | github

- Created a Python web application to monitor daily Bombay Stock Exchange prices by data scraping, data cleaning and offering **insightful stock performance analytics** via data visualization techniques**.**
- Facilitated seamless integration with machine learning algorithms for **advanced predictive analysis**