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, mathematics, data analytics

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, Computer vision, data mining

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

- **Accelerated development time by 30%** by collaborating in the design and development of cloud-based data applications, specializing in **machine learning models** for **predictive health management systems** in **aerospace equipment**.
- **Increased prediction accuracy by 25% and reduced downtime by 20%** by engineering and integrating 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**.
- **Improved system reliability by 40%** by working within a dynamic, **cross-functional** team to merge various technologies, ensuring the deployment of **cloud-hosted machine learning solutions** tailored for **real-time health monitoring.**

Machine Learning Student Researcher | UCI AI Center — Jun 2024 - Present

- **Improved model evaluation efficiency by 50%** by executing a comprehensive evaluation 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.
- **Increased model generalization accuracy by 15%** by analyzing model performance and generalization to external data - targeting **brain and lung cancers**, delivering insights into the effectiveness of foundational models in **clinical survival prediction**.

Natural Language Student Researcher | UCI INCHES Lab — May 2024 - Present

- **Reduced manual work** of **2 weeks to 1 day** by engineering a program in the **UCI INCHES Lab** with **Dr. Angela Lukowski** using Python and **spaCy** to automatically **parse narrative data** into propositional phrases.
- **Enhanced the efficiency by 70%** of analyzing **event memory studies** by implementing flexible rules to accommodate varied subjects and verbs within the narratives, ensuring the program's adaptability to different datasets.

Cloud Engineer | Airbus — Jul 2021 - Aug 2023

- **Facilitated the migration of 65%** of departments from **on-premises** to **cloud infrastructure** via collaborattion within a 10-member team to deliver secure and scalable cloud services on **AWS**,
- **Reduced monthly security alerts by 20%** by employing **analytical Bayesian methodologies** to detect and address potential security threats, ensuring **data integrity**.
- **Achieved annual cost savings of $70,000** by integrating a **recommendation engine** into production, utilizing historical usage patterns to optimize menu offerings.
- 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

- **Cut yearly expenses by $1 million** by engineering a **chatbot** via **natural language processing** and **Azure Bot Framework.**
- **Reduced customer service response time by 1.7 hours and achieved 30% surge in chatbot usage** via implementation of a **chatbot analyzer** capable of identifying areas of low performance and providing actionable insights for enhancement.
- **Increased positive feedback by 25%** by performing **A/B testing** on intents via feedback and incorporated them into language model.

PROJECTS:

Credit Reporting Consumer Complaints Analysis | github | medium

- Utilized Python with **Pandas** and **Seaborn** to process and visualize data from Equifax, Experian, and TransUnion, enhancing analytical precision and identifying key trends such as **credit score distributions**, **delinquency rates**, and **regional credit behavior patterns**.
- Developed predictive models using Python and NumPy to **forecast trends** and assess policy effectiveness

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.