

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; OpenCV; Seaborn; MilvusDB; Neo4J; Pandas; Keras; Numpy, spark, Hadoop, hive, mapreduce
- AWS - solutions architect, Open-source contributor of Mozilla - [github](#).

WORK EXPERIENCE:

Machine Learning Infrastructure Intern | **Safran**

Jul 2024 – Nov 2024

- Boosted equipment reliability to 97% by deploying machine learning models trained on vibrational frequencies using pytorch hosted in Docker containers, orchestrated via Kubernetes thereby scaling predictions.
- Accelerated equipment health prediction time by 18% using Python and Spark by parallelizing real-time aerospace equipment data workloads.

Machine Learning Student Researcher | **UCI AI Center**

Jun 2024 - Present

- Increased model efficiency by 11% for cancer survival prediction by modeling data pipeline that import datasets from TCGA, CPTAC and private cohorts with Python and PyTorch. Worked with Dr. Jana Lipkova.
- Increased model accuracy to 94% using python and pandas by analyzing model performance and generalization to external data delivering insights into the effectiveness of transfer learning in clinical survival prediction.

Natural Language Student Researcher | **UCI INCHES Lab**

May 2024 – Sep 2024

- Reduced manual work of 2 weeks to 1 day using Python and spaCy by automating parsing of narrative data into propositional phrases by engineering a Large Language model. Worked with Dr. Angela Lukowski.
- Enhanced the efficiency to 97% of analyzing event memory studies by implementing flexible rules to accommodate varied subjects and verbs within the narratives using spaCy tokens, ensuring the program's adaptability to different datasets.

Cloud Engineer | **Airbus**

Jul 2021 - Aug 2023

- Reduced monthly security alerts by 20% by employing analytical Bayesian methodologies using python and splunk to improve precision of alert systems by minimizing false positives and false negatives.
- Achieved annual cost savings of \$70,000 by integrating a recommendation engine built on python utilizing historical usage patterns to optimize cloud resource 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.

Natural Language Developer Intern | **Novartis**

Jan 2021 - Jun 2021

- Cut yearly expenses by \$1 million by engineering a chatbot on python via natural language processing and Azure replacing L1 customer support.
- Reduced customer service response time by 1.7 hours and achieved 33% surge in chatbot usage via a chatbot analyzer using python, NLP and pandas capable of identifying areas of low performance and providing insights for enhancement.
- Increased positive feedback by 23% by performing A/B testing on intents via feedback and incorporated them into language model.

PROJECTS:

Lane Detection for Autonomous Driving Using Attention | [github](#) | [medium](#)

- Designed and implemented a custom lane detection model using PyTorch and OpenCV with UNet-based architecture, incorporating residual blocks and attention mechanisms to enhance lane segmentation accuracy for autonomous vehicles.
- Developed image preprocessing and data augmentation pipelines, including edge detection, noise reduction, and morphological transformations, resulting in improved model performance on challenging lane detection scenarios.

Credit Reporting Consumer Complaints Analysis | [github](#) | [medium](#)

- Utilized Python with Pandas and Seaborn to process and visualize data from Equifax, Experian, and TransUnion, identifying key trends in credit score distributions using statistical modeling providing critical insights for strategic financial decision-making.
- Developed predictive models using Python and NumPy, applying regression analysis and time series forecasting to assess policy effectiveness and forecast financial trends, enhancing strategic planning.