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

- 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

- Data Structures; Database Management; Natural Language Processing; Programming, Computer vision, data mining

SKILLS:

- Python, C++, R, SQL; Pandas, Numpy; OpenCV, Matplotlib, Seaborn, Neo4J, spark, Hadoop, hive, mapreduce, predictive analytics, looker, Docker; Kubernetes
- AWS - solutions architect, Open-source contributor of Mozilla - github.

WORK EXPERIENCE:

Machine Learning Engineer Intern | **Safran** **Jul 2024 – Nov 2024**

- Saved $10,000 per aircraft per month by predicting equipment health with multiple models trained on vibrational, temperature, pressure, and usage cycle data using Python.
- Increased model accuracy to 96% by developing an ensemble model combining LSTM for temporal data and XGBoost for categorical features to improve accuracy and efficiency.

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

- Improved cancer survival prediction accuracy to 89% using transfer learning on patch image and whole slide images with PyTorch. Worked in collaboration with Dr. Jana Lipkova.
- Achieved 81% base model accuracy by applying Principal Component Analysis (PCA) to reduce dimensionality of extracted features to train SVMs, Linear Regression and KNN models.

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 narratives into prepositional phrases by engineering a Part Of Speech (POS) tagging model. Worked with Dr. Angela Lukowski.
- Achieved 97% efficiency in analyzing event memory studies by implementing flexible rules with tokens to handle varied subjects and verbs across datasets.

Data Scientist | **Airbus** **Jul 2021 - Aug 2023**

- Reduced monthly security alerts by 20% by employing analytical Bayesian methodologies using python to improve precision of alert systems by minimizing false positives and increasing recall.
- 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.
- Increased positive feedback by 23% by performing A/B testing on intents via feedback and incorporated them into language model and evaluated significance using t-tests.

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.