Pranav Gujarathi

2808 McKinney Ave Apt 718, Dallas, TX 75204 | gujarathi.pranav@gmail.com

LinkedIn: linkedin.com/in/pranav-gujarathi/ | **GitHub:** github.com/pranavdg1997 | website

Education

Master of Science in Data Science | Indiana University, Bloomington |GPA: 3.54 **August 2019** - **May 2021**

Bachelor of Technology in Electrical Engineering | Indian Institute of Technology, Roorkee **July 2014** - **May 2018**

Work Experience

Data Scientist | Walmart Global Tech (Dallas, TX) **January 2022 – Present**

- Deployed and managed large-scale anomaly detection engine with into production with real-time user feedback loop.
- Developing automated framework for causal interpretation of anomalies and derive appropriate response action.

Senior Data Science Associate | ZS Associates (Los Angeles, CA) **May 2021 – December 2021**

- Deployed a product with favorable client feedback and improved performance in the form of a cross platform application.
- As part of the project, utilized Python libraries, Deep Learning frameworks and transformer models to implement a Natural Language Inference pipeline, i.e., extracting domain-relevant inferences from textual data (news articles, publications, etc.).
- Spearheaded the Data Science team for successful discussions with client teams and internal stakeholders.

Graduate Researcher | **Indiana University (Bloomington, IN)** **January 2020 – May 2021**

- Building a Mind Lab: Designed and implemented pipelines to successfully conduct experiments as part of NSF funded project under the guidance of Professor Justin Wood. The project involved working across topics in **Computer Vision** and **Deep Reinforcement Learning**.
- IUPUI Data Lab**:** Conducted research and experiments in Natural Language Processing models and architectures towards a successful end to end process from ideation to eventual publication under the guidance of Prof Sunandan Chakraborty.
- Kelley School of Business: Successfully deployed an MLOps pipeline starting from a PoC formulation to a GUI dashboard using Big Data libraries and cloud-based parallel computation.

Data Science Associate | **ZS Associates (Pune, India)** **July 2018 - July 2019**

- Designed a novel ML based solution for marketing strategy planning which significantly improved revenue compared to legacy methods while minimizing spending.
- The solution utilized multivariate Time Series forecasting and Linear Optimization, implemented in Spark (PySpark).
- Successfully pitched the pipeline in **client facing** interactions and worked on end-to-end development of the pipeline from POC to deployment, taking complete ownership of the data science aspect of the project.

Technical Skills

Programming Languages – Python (PyTorch, Tensorflow, NumPy, Pandas, GPU use, OpenCV, Scikit learn, Jupyter), Swift, R (data.table, h2o, RShiny), C/C++ (Data structures and algorithms), BI Platforms (Tableau, Excel), Big Data(SQL, Spark), Version Control (Git, bash, SVN), Cloud computing (AWS, Databricks), Front-end and Web development (Flask, Dash, C#), Git

Conceptual skills– Statistical Experience and Quantitative analysis, Data Mining, Predictive modelling**,** Programming fundamentals and algorithm development**,** Computer Vision, Natural Language Processing (NLP), Reinforcement Learning, Information Retrieval, Optimization algorithms, Regression, Classification, Clustering, Time series forecasting, Recommendation systems.

Projects

Chick AI: Understanding Animal behavior using Computer Vision and Reinforcement Learning **January - April 2021**

- Experimented with various computer vision architectures to simulate biological object detection process in baby chickens' brain to a digital environment.
- Achieved improved accuracy (76% to 93%) as well as computation cost (-34%) for object detection using models such as SIMCLR, Autoencoders and as well as improvement in RL tasks using state-of-art models (A2C and PPO models) as part of experiments to reverse engineer visual understanding in animal brains.

Cause-Effect Entity Recognition using Natural Language Processing **May 2020 – May 2021**

- Achieved state-of-art performance (91% Recall) for the NLP task of predicting causal-inference based entity recognition from text, by implementing modified Transformer models in PyTorch.

Navigation System for Self-Driving Vehicles(read more) **January 2021 – May 2021**

- Successfully built a self-driving bot in a simulated environment (Duckietown) removed the need of labelled data for training by leveraging Deep Reinforcement Learning based model.
- Improved the performance of the system to over 150% using Deep Deterministic Policy Gradients (DDPG) algorithm, and finally to 180% by using a modified version of the AlphaGo algorithm.

High on Data **- Data Science Tool to Tackle Opioid Crisis** | AT&T FirstNet Hackathon **September 2019**

- Built a protype AI tool to scrape, process and predict opioid addiction from Twitter using Natural Language Processing.
- Wrapped the tool into an easy-to-use Visualizations and GUI to create better experience for end users.

Publications, Achievements and extra-curricular

- Published **"Controlled-rearing studies of newborn chicks and deep neural networks"** at **Shared Visual Representations in Human & Machine Intelligence workshop, NeuRIPS 2021** winning best paper award at the event.
- Published **"Using Causality to Mine Sjögren's Syndrome related Factors from Medical Literature"** at ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)
- Awarded ***Luddy Outstanding Research Award*** for research contributions during MS degree. (May 2021)