Pranav Gujarathi

3210 K, John Hinkle Place, Bloomington, IN 47408 | (812) 803-8003 | pgujarat@iu.edu
linkedin.com/in/pranav-gujarathi/ | github.com/pranavdg1997 | website

Education

Master of Science in Data Science | Indiana University, Bloomington |GPA: 3.52 **August 2019** - **May 2021**
Relevant Coursework: *Deep Learning Systems, Advanced Natural Language Processing, Reinforcement Learning*
Bachelor of Technology in Electrical Engineering | Indian Institute of Technology, Roorkee **July 2014** - **May 2018**

Work Experience

NLP Research Assistant | Indiana University (Prof Sunandan Chakraborty)| Website **May 2020 – Current**

- Actively working on an NSF-funded project, '*Capturing Dynamism in Causal Relationships: A New Paradigm for Relationship Extraction from Text*', which involves extracting entities and latent causal relationships from domain-specific text.
- Used Transformer architectures and transfer learning from state-of-art models to achieve 89% F1 score in detecting cause-effect relationships.
- To counter lack of labelled data, currently exploring Deep Reinforcement Learning algorithms like Soft Actor Critic and Proximal Policy Optimization for unsupervised learning and entity extraction. Refer this Letter.

Associate Instructor | Indiana University (Prof Jingwen Yan) **May 2020 – November 2020**

- Preparing lecture slides and Pytorch based coding assignments for "Application of AI in biomedical informatics"

Graduate Research Assistant | Kelley School of Business, Indiana University (Prof Jeff McMullin) **January 2020 – May 2020**

- Used Big Data tools and parallelization to reduce data preprocessing workflows to less than 10% of original time.
- Successfully developed an interactive dashboard (GUI) web app to process data with user input and visualize network data.

Data Science Associate | ZS Associates **July 2018 - July 2019**

- Implemented numerous descriptive and predictive Data Science tasks on healthcare datasets, and upgrading conventional pipelines in classification, regression and forecasting to state-of-art models.
- Designed a novel solution which utilized combination of multivariate Time Series forecast and Linear Optimization to formulate optimal marketing strategy for the client which can significantly improve profits.
- Successfully pitched the pipeline in client facing interactions and worked on complete development of the pipeline from POC level to PySpark based deployment in production, followed by further improvements after feedback .

Data Science Summer Intern | Nokia Solutions and Networks **May 2017 – August 2017**

- Developed an automated system for network fault and diagnosis using semi-supervised Machine Learning.
- Improved accessibility of the system by building a Python based GUI, converting the code into an end-to-end solution.

Technical Skills

Programming Languages – Python (Pytorch, Tensorflow, Keras, 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, PySpark), Version Control(Git, bash, SVN), OO/Software design, Front end and Web development(Node.js, Flask, Dash, C#), Unix/ Linux
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, Model interpretation techniques, Robotics fundamentals.

Projects

Navigation System for Self-Driving Vehicles(read more) **January – April 2020**

- 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

Multi-Modal Yelp recommendation system (read more) **August-December 2019**

- Built a solution which analyzes historical Yelp reviews for customer-personalized restaurant recommendations.
- Optimized the system for fast computation using PySpark based processing on parallel nodes and Pytorch for GPU, with pipelines to analyze numerical, textual as well as image data.
- Incorporated additional utility of deeper business insight by using model interpretability algorithms.

***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

***RecruitZ* - Smart Recruitment Solution for Companies** | ZS QUEST Hackathon **February 2019**

- Designed and contributed to a novel Application Tracking System (ATS) project in a 36-hour hackathon.
- Used Information retrieval combined with Natural Language embeddings to significantly improve over existing solutions.

Achievements and extra-curricular

- Awarded the ***Best use of MPH Open Data*** for the '*High on Data*' project at AT&T Firstnet Hackathon. (February 2019)
- Awarded the ***Best User Interface*** and ***Best Return on Investment*** for the '*Recruitz*' project. (September 2019)
- Among the top 2% candidates selected for Data Science Associate position at ZS via a hackathon (October 2017).
- Served as writer and finance coordinator for campus news agency (September 2014 - May 2018)
- Volunteered for disaster relief for National Service Scheme(Government of India) (September 2014 - May 2018)