

ABHIBHA GUPTA

abg96@pitt.edu ♦ +1 4129547510 ♦ abhibha1807.github.io ♦ linkedin.com/in/abhibha-gupta/

EDUCATION

University of Pittsburgh August 2022 - June 2024 (Expected)
MS in Information Science GPA: 4/4

Indian Institute of Information Technology, Nagpur July 2017 - June 2021
BTech in Computer Science and Engineering GPA: 3.45/4

TECHNICAL SKILLS

Languages: Proficient - Python, R Intermediate - MySQL, C++, C, Bash **Data Processing:** Python - Pandas, NLTK, CoreNLP, Gensim R - Tidyverse, igraph **ML/DL:** Python - PyTorch, Tensorflow, Fastai, Keras, scikit-learn, spaCy, Hugging Face, Langchain R - Caret, glmnet
Visualisation: Python - Matplotlib, Seaborn R - ggplot

Coursework - Machine Learning in R (Code), Databases, Algorithms, Data Mining, Artificial Intelligence, Information Storage and Retrieval, Natural Language Processing, Deep Learning, Computer Vision, Neuro-Fuzzy Techniques, Probability, Graph Theory, Data Science, Bio-informatics

PUBLICATIONS

- **Argumentative Stance Prediction: An Exploratory Study on Multimodality and Few-Shot Learning - 2023**
Arushi Sharma, [Abhibha Gupta](#), Maneesh Bilalpur- Empirical Methods in Natural Language Processing (EMNLP)
- **Enhancing Visual Perception in Novel Environments via Incremental Data Augmentation Based on Style Transfer - 2023** [Paper](#)
[Abhibha Gupta](#), Mansur Maturidi Arief, Rully Agus Hendrawan- International Conference on Robotics and Automation (ICRA) *Under Review*
- **Towards Accurate and Clinically Meaningful Summarization of Electronic Health Record Notes: A Guided Approach - 2023** [Slides](#)
Zhimeng Luo, Yuelu Ji, [Abhibha Gupta](#), Zhuochun Li, Adam Frisch, Daqing He, IEEE International Conference on Biomedical and Health Informatics
- **Neural architecture search for pneumonia diagnosis from chest X-rays. - 2022** [Paper](#), [Article](#)
[Abhibha Gupta](#), Parth Sheth, Pengtao Xie, Journal of Nature Scientific Reports.
- **Disambiguating spatial prepositions: The case of geo-spatial sense detection. - 2022** [Paper](#)
Mansi Radke, [Abhibha Gupta](#), Kristin Stock, CB Jones, Transactions in GIS Journal

RESEARCH EXPERIENCE

Guided summarization of clinical notes - University of Pittsburgh (IRIS Lab) 📍 Pittsburgh, US
Graduate Student Researcher, Advisor: Daqing He 📅 May 2023 – Present

- Collaborated with medical practitioners to develop a comprehensive summary template consisting of important medical entities covering patient demographics, chief complaint, OPQRST assessment, diagnostics, treatment, etc.
- Fine-tuned a BART-based guided summarization model with textual guidance extracted from a sentence classifier trained to determine the importance of sentences within clinical notes.
- Implemented a fact-checking tool for validating the factual correctness of predicted summaries against the ground truth by fine tuning a Bio-ClinicalBERT-based Named Entity Recognition (NER) model
- Experimenting with medically focused prompt engineering techniques by running multi-GPU inference using Huggingface Accelerate on large language models (LLMs) like LLaMA and Alpaca for task of "faithful" clinical note summarization. Reduced inference time by **41%**.

Improving rare traffic sign recognition via data augmentation - Stanford University (SISL Lab) 📍 Remote
Independent Student Researcher 📅 May 2023 – September 2023

- Proposed a novel approach to enhance visual perception in autonomous agents by tackling the uncharted territory of "unknown unknowns" in real-world scenarios, particularly focusing on degraded signs not seen during training.
- Leveraged the Variational Prototyping Encoder (VPE) combined with neural style transfer to incrementally augment data, resulting in a significant boost in model robustness when faced with novel inputs.
- Demonstrated the effectiveness of the proposed methodology on the German Traffic Sign Recognition Benchmark (GTSRB) dataset, **achieving an F1 score of 86%** (36% increment over baseline) , underscoring the model's adaptability and resilience to previously unseen data variations.[Code](#)

Neural Architecture Search (NAS) for Pneumonia Diagnosis - University of California, San Diego 📍 Remote
Research Intern, Advisor: Pengtao Xie 📅 May 2021 – April 2022

- Implemented Neural Architecture Search (NAS) for improving pneumonia detection from Chest X-Ray images by leveraging the "Learning By Teaching" (LBT) framework. The approach inspired by the teacher-student learning paradigm involves a teacher model guiding a student model, emphasizing continuous learning and improvement.
- The LBT enhanced NAS method surpasses previous methods like [DARTS](#) and [PC-DARTS](#), achieving a **5.1% increase in accuracy**. The optimal LBT-based NAS, specifically refined for pneumonia detection, achieves a high **97.6% ROC-AUC score** and is **4% more compact** than DARTS. [Code](#)
- **Reading by Translating:** Implemented the 'Reading by Translating' framework that improves the task of 'Machine Reading' i.e extracting meaningful instances from the dataset. Involves 2 transformer based encoder-decoder models, that are trained mutually on the task of Machine Translation to learn importance weights assigned to the dataset instances.

PROJECTS

Predicting optimal coating constituents for reducing corrosion - PPG Paints

Conducted exploratory data analysis and feature selection in R with Ggplot library. Employed Bayesian models (splines, XGBoost, Random Forests, MARS, Neural networks) to predict optimal constituents, successfully reducing the fraction of corroded area to **less than 33%**.[Code](#)