Akshit Achara

@ acharaakshit@gmail.com | LinkedIn | GitHub | Portfolio | London, UK

Education

Birla Institute of Technology and Science — Pilani, India
M.Sc. in Mathematics — *Aug 2016 – Jul 2020*

- Ranked best private university in India; Awarded highest possible grade in Master Thesis

General Electric — Bangalore, India
Edison Engineering Development Program — *Oct 2020 – Sept 2022*

- Highly selective program that accelerates participants' professional and technical development within GE's Advance Courses and a variety of business-critical engineering assignments.

King's College London — London, United Kingdom
PhD in AI for Medical Imaging — *Oct 2024 –*

- Conducting rigorous machine learning research for fairness in deep learning models for medical imaging

Publications/Pre-Prints/Patents

- **Achara, Akshit**, Sanand Sasidharan and Gagan N. "Efficient Biomedical Entity Linking: Clinical Text Standardization with Low-Resource Techniques." BioNLP Workshop @ ACL (2024)
- **Achara, Akshit**, and Ram Krishna Pandey. "Revealing the Underlying Patterns: Investigating Dataset Similarity, Performance, and Generalization." Neurocomputing (2024) 127205.
- Pandey, Ram Krishna, and **Akshit Achara**. "TrueDeep: A systematic approach of crack detection with less data." Expert Systems with Applications (2023) 122785.
- Pandey, Ram Krishna, and **Akshit Achara**. "CoreDeep: Improving Crack Detection Algorithms Using Width Stochasticity." arXiv preprint arXiv:2209.04648 (2022). (Modified version accepted for an oral presentation at *CVIP*)
- **Achara, Akshit**, Anuradha Kanamarlapudi, Sai Vivek Velkuru and Sanand Sasidharan. "Patient Centric Generative PROs." US Patent 90289027 (Application filed). (2024)
- Soumya Ghose, Sanand Sasidharan, Sanghee Cho, **Akshit Achara**, Sundarajan Mani, Ravi Bharadwaj, Brion Sarachan, Rakesh Mullick, Anuradha Kanamarlapudi and Fiona Ginty. "Natural Language Based Clinical Guidance System for Care Pathway." US Patent 90281596 (Application filed). (2024)
- Sasidharan, Sanand, **Akshit Achara**, Anuradha Kanamarlapudi, Shivappa Goravar and Kshireswar Bhoi. "AI-Assisted EMR Workflows." US Patent 90289229 (Application filed). (2023)

Research Experience

Explainable AI
King's College London — *Oct 2024 –*

- Axiomatic Analysis of Explainable AI approaches like Integrated Gradients, Baseline Shapley (BShap), DeepLIFT, LRP, gradCAM, guided gradCAM, etc. The goal is to apply explainable AI approaches that satisfy all the relevant axioms to medical imaging.
- Studying the mode collapse for language embedding based approaches to find potential for retrieval performance.

Practical NLP for Oncology
GE Research — *Apr 2021 – Sept 2023*

- Developed a combined annotation independent token classification and entity linking approach on patient reports by leveraging structures of medical ontologies like SNOMED and UMLS.
- Attention Analysis of encoder based models for understanding close and far off relationships between medical terms to improve the granularity of phrases. For example, "Metastatic Lung cancer" is much more informative compared to "Lung cancer" but "Metastatic" might be mentioned in a far away context.

- Interaction with top oncologists from NCCS, Singapore to gain the clinician feedback on report summarization. Discussed challenges like "poorly differentiated carcinoma" being predicted as a negation by assertion models and their potential solutions.
- Applied FedAvg method to perform federated learning on real-world hospital data and observed the practical relevance of the approach despite multiple works proving the drawbacks like client drift theoretically.

Knowledge-Driven Deep Learning

GE Research *Apr 2021 – Sept 2023*

- Extensive empirical experimentation to understand the behaviour and generalization of a breadth of models on image segmentation. Trained multiple models on a primary dataset and evaluated generalization performance of several similar and different secondary datasets.
- Developed techniques to select a coreset from an overall dataset. The model performance on coreset outperformed state-of-the-art models on crack detection.
- Used stochasticity in width of cracks as a basis for data augmentation resulting in high quality segmentation masks.
- Developed a web-based MRI viewer with a brain MRI segmentation model utilizing multiple augmentation methods and sampling. We achieved a test F-score of 0.84 on real-world data and the overlayed segmentation masks were well accepted by the clinicians.

Benchmarking miRNA Target Prediction Methods

Master Thesis, Technical University of Munich *Aug 2019 – Dec 2019*

- Developed mirtargetbenchmark: A machine learning tool to benchmark miRNA target prediction tools.
- Applied multiple data preprocessing techniques on noisy expression data before computing the gene-miRNA correlation to prepare ground truth.
- Applied multiple regularization techniques like Ridge and Lasso regression to avoid overfitting of the linear models.
- Approaches used the number of binding sites and confidence scores of sequence based target prediction tools as predictions for benchmarking against ground truth correlation between gene and miRNA expression.

Constraint Optimization in R

Student Developer, Google Summer of Code *Jun 2020 – Oct 2020*

- Development of rminizinc: An interface to MiniZinc in R as a part of GSoC 2020.
- Directly worked with the creators of MiniZinc to understand the the language better and designed the framework accordingly.
- Worked on advanced R programming to create custom classes for corresponding structures and classes in MiniZinc.
- Integrated existing MiniZinc solver classes by Rcpp to allow solving problems with multliple solver backends like GeCode, Chuffed, etc.
- The package allows solution of satisfaction, minimization and maximzation and contains solutions to a variety of constraint optimization problems as examples.

Classification of Respiratory Patterns for Dynamically Adjustable X-Rays

Researcher, Machine Vision lab, CSIR-CEERI *Jan 2019 – May 2019*

- Developed a deep learning model for classification of breathing patterns in multiple patients to avoid the harm to healthy cells because of X-rays.
- Manually collected volunteer respiration data using kinect RGB-D camera. Applied noise reduction and augmentation on frames for data pre-processing.
- Developed a VGG-16 based classification model initialized with imagenet weights resulting in a test accuracy of 86.5%.

Work Experience

GE Healthcare Research — Bangalore, India

Research Engineer *Oct 2023 – July 2024*

- Development of a time-aware clinical trial matching application utilizing an ensemble extractive and generative NLP techniques to recommend personalized clinical trials for patients.

GE Research — Bangalore, India

Edison Engineer, Research Engineer *Apr 2021 – Sept 2023*

- Finetuning and Reinforcement learning of a GPT based generative text model on customer-engineer interactions to provide automatic solutions to customer queries.

- Developed an automatic Clinical Trial Matching tool including techniques to convert unstructured Inclusion-Exclusion Criteria into structured numerical, ordinal and nominal fields leveraging named entity recognition,entity linking and Regex based approaches.
- Timeline-based summarization of patient reports for Oncology using transformer based pipelines coupled with a multitude of rules created in collaboration with top oncologists.
- Developed a cross-silo federated learning framework for token classification on real-world patient data with multiple collaborators to enable distributed training on image segmentation.
- Developed a docker-based end-to-end Micro-grid Energy Management System, deployed at multiple production sites.

GE Healthcare Bangalore, India

Edison Engineer *Oct 2020 – Mar 2021*

- Developed and Integrated a Voice to Action converter into a web based diagnostic viewer for radiologists to reduce the time consumed while traversing between multiple screens and exams.

GE Healthcare Bangalore, India

Intern *Summer 2018, 2019*

- Developed a multi-threading approach in C for asynchronous logging to improve the performance of MRI Scanners.
- Designed and developed a cloud based application to handle real-time multi-modal data and real-time alert systems including various algorithms like Triple Low and Anesthetic agent check to determine the mortality rate of a patient.

Skills

Recent Courses: Measure-Theoretic Probability, Fundamental Theory of Statistical Inference

Programming: C, C++, Java, Python, JavaScript, R, SQL, Git, Docker

Libraries: Scikit-Learn, PyTorch, Keras, TensorFlow, Pytorch-Lightning

Languages: Hindi (Native), English (Professional), Sanskrit (Novice)

Other Activities

- **Academic Services:** Reviewer (NeurIPS 2024, ICLR 2025)
- Won the 1st Place Award in the Apogee Innovation Challenge at BITS Pilani that resulted in a summer internship at GE Healthcare.
- Served as a Joint Secretary for the Mathematics Association at BITS Pilani conducting various contests with 200+ participants.
- Won 3rd Place in the football tournament comprising of 20 teams at GE, Bangalore, India.
- Won the One Dream Corporate Cricket Cup which is an open tournament in Bangalore, India comprising of 20 teams.