Poornash Anandan Sangeetha

(A.S. Poornash)

M.S., Machine Learning
Carnegie Mellon University, Pittsburgh, PA

Email: *aspoornash2355@gmail.com*
Google Scholar: *Scholar Profile*
Webpage: *poornash.github.io*
GitHub: *github @ poornash*

EDUCATION

- **Carnegie Mellon University (CMU)** *Pittsburgh, USA*
  Masters of Science (M.S.) in Machine Learning (2025 - 2027) *GPA: **Yet to give exams***
- **Indian Institute of Technology (IIT), Patna** *Bihar, India*
  Bachelor of Technology in Computer Science and Engineering (2021 - 2025) *GPA: **9.71/10.00***
  – Ranked first (**Institute Rank 1 & President's Gold Medalist**) among **550** students and in the '25 batch at IIT Patna.

PUBLICATIONS

- **ToxVidLM: A Multimodal Framework for Toxicity Detection in Code-Mixed Videos**   **ACL Findings 2024**
  Authors - A.S. Poornash, Krishanu Maity*, Sriparna Saha, Pushpak Bhattacharyya*
  – Accepted at Findings of the Annual Conference of the Association for Computational Linguistics, 2024

- **Seeing Beyond Words: Multimodal Aspect-Level Complaint Detection in Ecommerce Videos**   **ACM Multimedia 2024**
  Authors - A.S. Poornash, Rishikesh Devanathan*, Apoorva Singh*, Sriparna Saha*
  – Accepted at the ACM International Conference on Multimedia as an *Oral Presentation* (ACM MM, 2024)

- **A Stochastic Optimization Framework for Private and Fair Learning from Decentralized Data**   **PPAI Workshop, AAAI 2025**
  Authors - Devansh Gupta, A.S. Poornash, Andrew Lowy, Meisam Razaviyayn
  – First version is accepted at the PPAI Workshop, AAAI 2025 and the second is under review at a reputed ML conference.

- **APTSumm at BioLaySumm Task 1: Biomedical Breakdown, Improving Readability by Relevancy-Based Selection**   **BioNLP Workshop, ACL 2023**
  Authors - A.S. Poornash, Atharva Deshmukh, Archit Sharma, Sriparna Saha
  – Accepted at the BioNLP Workshop, Association for Computational Linguistics, 2023

- **ToxVI: A Multimodal LLM-based Framework for Generating Interventions in Toxic Code-Mixed Videos**   **CIKM 2024**
  Authors - A.S. Poornash, Krishanu Maity*, Sriparna Saha, Kitsuchart Pasupa*
  – Accepted at the ACM International Conference on Information and Knowledge Management, 2024

- **HateThaiSent: Sentiment-Aided Hate Speech Detection in Thai Language**   **IEEE Transactions (TCSS)**
  Authors - A.S. Poornash, Krishanu Maity*, S Bhattacharya, S Phosit, S Saha, Kitsuchart Pasupa*
  – Accepted in the IEEE Transactions on Computational Social Systems Journal, 2024   (* *= Equal Contribution*)

SCHOLASTIC ACHIEVEMENTS

- **Academics:** IIT Patna 2021 - 2025 undergrad **Batch Topper (Institute Rank 1)** with a **GPA** of **9.71/10.00**
- **Amazon ML Challenge '24:** Secured **6th** position among **75,000+** participants in an ML-based hackathon (Certificate)
- **IUSSTF-Viterbi India Scholar '24:** One of only **15** selected student awardees from India to conduct research at the University of Southern California (**USC**), by Department of Science & Technology (DST), Govt of India (Certificate)
- **MITACS GRI Research Fellowship '24:** Awarded to conduct research at University of Alberta, Canada (Award Letter)
- **DAAD-WISE Research Fellowship '24:** Awarded to conduct research at TU Darmstadt, Germany (Award Letter)
- **Inter IIT Tech Meet 11.0 '23:** Bagged **1st position** among all **IITs** in an Natural Language Processing-based high-prep event DevRev - Expert Answers in a Flash: Improving Domain-Specific Question Answering. (Repository link)
- **SPARK Fellowship '23:** Selected for the prestigious summer research program of **IIT Roorkee** in my sophomore year.
- **KVPY-SX Fellow '21:** Research Fellowship awarded by DST, Govt of India, All India Rank **517** out of **160,000** students
- **JEE Advanced '21:** Secured a **99.05 percentile** score (All India Rank - **2355**) out of **250,000** candidates nationwide.
- **JEE Mains '21:** Secured a **99.79 percentile** score out of **1.1 million** candidates from across the country.

TECHNICAL SKILLS

- **Programming Languages:** Python, MATLAB, Prolog, Lua, C, C++, Java, Kotlin, Verilog, SQL.
- **Frameworks & Libraries:** PyTorch, TensorFlow, Keras, Caffe, NumPy, Pandas, SciPy, Scikit learn, Hugging Face Transformers / Accelerate / PEFT, LangChain, LlamaIndex, vLLM, Haystack, Ollama, LangSmith, Neo4j, NLTK, spaCy, TextBlob, Gensim, TensorFlow Federated, Opacus, FederatedAI/FATE, Fairlearn, ITK, OpenCV, Scikit-Image, PIL, Napari, MONAI, Librosa.
- **Development Tools:** AWS, Git, Anaconda, Android Studio, Arduino, MySQL, Firebase, Firestore, Docker, LaTeX.
- **Web Development:** HTML, CSS, JavaScript, PHP.

INDUSTRY & RESEARCH EXPERIENCE

- **Automated HITL Review for Cross-Border Product Benchmarking @ (Amazon Applied Science)** *Jan. 2025 - July. 2025*
  Manager: Ramya Toshniwal | Position: ML Intern @ Amazon Applied Science
  - Developed and implemented an end-to-end automated workflow using vision-language models to search and classify for Amazon and competitor product pairs (across multilingual and multimodal contexts) into nuanced match categories, reducing manual review time by over **90%** while increasing consistency and scalability.
  - Designed a robust, SOP-driven agentic loop combining image and text analysis, resulting in a significant improvement in match classification accuracy (F1 score up to **0.87**) and optimized resource usage, as validated against months of ground truth datasets.

- **SteFFLe: A Stochastic Framework for Private, Fair, and Federated Learning @ (USC Viterbi)** *May. 2024 - July. 2024*
  Guide: Dr. Meisam Razaviyayn | Lab: University of Southern California, ODDS Research Group | Venue | | Paper | | Code |
  - Contributed to developing a Stochastic, Differentially Private, Fair, and Federated framework with proven inter-silo record-level differential privacy **(ISRL-DP)** guarantees. Accepted at the **PPAI Workshop** of **AAAI 2025**.
  - Proved algorithm convergence in heterogeneous data distributions across silos and for non-binary attributes.
  - Designed experiments and made privacy-utility tradeoff curves for various (ϵ, δ) privacy budgets and heterogeneity levels in federated settings, outperforming **two SOTA** approaches with a **75.47%** and **52.93%** reduction in demographic parity violations.

- **Multimodal LLMs for Toxicity Detection and Intervention in Code-Mixed Posts** *Nov. 2023 - April. 2024*
  Guide: Dr. Sriparna Saha | Lab: IIT Patna, AI-NLP-ML Group **ACL:** | Paper | | Code | **CIKM:** | Paper | | Code |
  - Developed ToxVidLM, a multimodal multitask framework for code-mixed video toxicity analysis, achieving **94.29%** accuracy and **94.35%** F1-scores on the toxicity classification task, accepted to **ACL Findings 2024**.
  - Designed a low-compute modality alignment technique that integrates visual and acoustic data with text using soft multimodal tokens generated with large language models' embedding matrices, improving primary task's F1 scores by approximately **6.43%**.
  - Extended this work to generate interventions for enhancing the explainability of predictions, which was accepted to **CIKM 2024**.

- **Lay Summary Generation for Biomedical Texts to Enhance Readability** *Feb. 2023 - July. 2023*
  Guide: Dr. Sriparna Saha | Lab: IIT Patna, AI-NLP-ML Group | Paper | | Leaderboard |
  - Developed a three-step abstractive summarization framework for biomedical documents to generate lay summaries, accepted via the BioLaySum Shared Task to the **BioNLP Workshop** at **ACL 2023**.
  - Our team **APTSumm** achieved the **second-highest** readability scores (**FKGL**: 12.22, **DCRS**: 8.99) in the BioLaySum shared task.
  - Obtained strong performance in relevancy metrics (ROUGE-1: 48.32, ROUGE-2: 14.91, ROUGE-L: 45.41) using a contrastive learning-based RoBERTa re-ranking strategy, enhancing summary quality through cosine similarity matching and beam search.

- **Sentiment-Aided Hate Speech Detection in Low-Resource Languages** *April. 2023 - April. 2024*
  Guide: Dr. Kitsuchart Pasupa | King Mongkut's Institute of Technology Ladkrabang, Thailand | Paper | | Code |
  - Developed a SOTA **lightweight** multi-task framework combining FastText and mBERT embeddings with capsule networks, advancing hate speech detection in Thai, published in the **IEEE Transactions** in Computational Social Systems **Journal** (*IF* - 4.5).
  - Achieved 89.67% accuracy and 89.79% macro-F1 score for hate speech; 80.92% accuracy and 80.97% macro-F1 score for sentiment analysis, outperforming **three SOTA** approaches by approximately **3 - 5%** with **10x** lesser parameters.

- **Median Nerve Detection and Segmentation for Carpal Tunnel Syndrome Diagnosis** *May. 2023 - July. 2023*
  Guide: Dr. Phaneendra K. Yalavarthy | Lab: IISc Bangalore, Medical Imaging Group (MIG) | Presentation | | Certificate |
  - Worked on medical image segmentation and classification tasks for Carpal Tunnel Syndrome diagnosis using **Ultrasound** videos.
  - Experimented on a dataset of 99 patients, achieving **Dice scores** up to **73.53%** and **IoU** up to **68.4%** for MN segmentation.
  - Led development and systematic experiments with **YOLO V3** and **U-Net** models, improving segmentation accuracy notably and serving as a lightweight alternative for real-time scenarios.

- **Expert Answers in a Flash: Domain-Specific Information Retrieval and Question Answering** *Jan. 2023 – Mar. 2023*
  Team project at Inter-IIT Tech Meet 11.0 - Bagged 1st place among all IITs in India | Code |
  - Developed a multi-stage Information Retrieval QA pipeline using **BM25 retriever** (95.4% top-5 recall, 16ms/query), transformer-based **cross-encoders**, and **generators** for domain-specific question answering.
  - Leveraged **knowledge distillation** (MiniLM-L-12-v2 → L-2-v2) and theme-specific re-ranking, achieving 89.3% top-1 ranking accuracy and 89.8% overall QA accuracy. Engineered a novel **answerability classifier** (perceptron) aggregating retriever, ranker, and reader signals, reaching 97.6% accuracy, 98.5% F1, and **sub-1s** end-to-end latency.

POSITIONS OF RESPONSIBILITY AND EXTRA-CURRICULAR ACTIVITIES

- **NJACK Machine Learning Division (Computer Science club of IIT Patna)** *2022 - 2024*
  Lead IIT Patna
  - Organized a Machine Learning **hackathon** in collaboration with the startup Chi SquareX during the Tech fest of IIT Patna.
  - Conducted **talks** by industry professionals, researchers (alumni), **research paper reading**, and **educational sessions** regularly.

- **Hexachrome (Puzzle Club of IIT Patna)** *2022 - 2024*
  Co-Founder & Sub-Lead IIT Patna
  - **Conducted** various Speedcubing **competitions** with 2x2, 3x3, 4x4, Pyraminx, etc. Designed and released weekly Sudoku puzzles.