Priyanshu Dhingra

priyanshudhingra1@gmail.com | +91-9729823428
Website | Google Scholar | Github | Linkedin

Objective

Dedicated researcher passionate about understanding intelligence in AI. Seeking research opportunities to deepen my knowledge of natural language processing and language model foundations. Strong academic foundation, a commitment to research excellence, and a desire to contribute meaningfully to the scientific community.

Education

Rajiv Gandhi Institute of Petroleum Technology Dec 2021 - May 2025
Bachelor of Technology - Information Technology ***CGPA: 9.36/10***
Relevant Coursework: Artificial Intelligence, Data Mining, Deep Learning, Computer Vision and Pattern Recognition, Digital Image Processing, Object Oriented Programming, Database Management Systems, Discrete Maths, Data Structures and Algorithms, Design and Analysis of Algorithms, Operating Systems, Computer Networks, Linear Algebra and Complex Analysis, Statistical Methods and Data Analysis, Real Analysis and Calculus

Work and Research Experience

Finastra Software Solutions, Bangalore Jan 2025 - Present

Associate Software Engineer - Data Science and AI

Project: Kondor Assistant

- Automated migration of 1,000+ QA cases into JSON test sets with LangChain + Azure OpenAI, reducing manual effort by 95%.
- Increased LLM response accuracy to 96% by building a scalable test automation framework, debugging failing cases, and strengthening evaluation logic.
- Implemented comprehensive data validation framework including SQL sanity checks, schema coverage analysis, and vector store enhancements with contextual Q&A pairs, strengthening governance and query reliability
- Built automated pipelines for accuracy/cost reporting and advanced SQL parsing for table-column relationship mapping.
- Diagnosed and resolved 15+ production issues including hyperlink errors, partial keyword failures, and ordering mismatches, boosting operational stability.

Carnegie Mellon University Sept 2024 - Jan 2025

Research Intern

Project: Vision transformer for 3D Subtomogram Alignment

- Co-authored a paper accepted at IEEE/CVF CVPR 2025, contributed to literature review, model design, experiments, and reproducible PyTorch/CUDA codebase.
- Helped design and implement components of the Polyshift and MARE modules to improve shift/rotation equivariance in ViT-based 3D alignment.
- Built pipelines for synthetic 3D subtomogram generation and 2D-slice visualizations used for dataset construction, debugging, and experimental evaluation.
- Implemented a Fourier-domain missing-wedge mask + interpolation pipeline as an exploratory approach to reduce anisotropic artifacts and improve alignment robustness.
- Debugged and optimized BOE-ViT training, enabling large-batch scalability and full reproducibility of results.
- Contributed to experiments demonstrating 77% lower rotation error and 62% lower translation error at low SNR (0.01), outperforming prior methods.

Nanyang Technological University Mar 2024 - Aug 2024

Research Intern

Project: Automatic Speech De-identification Systems

- First-author publications in three IEEE conferences (IALP 2024, TENCON 2024, ICAICTA 2024) and 1 journal (IJALP 2024) on speech privacy and de-identification.
- Developed an automatic speech de-identification system, achieving 0.79 timestamp-aligned F1 score for detecting and masking PII in audio.
- Innovated an LLM-based synthetic data generation pipeline (similar/diverse augmentation), paired with WhisperSpeech TTS to create speech–text pairs and address PII data scarcity.
- Designed and validated an Entity-Aware ASR model with joint ASR–NER training, outperforming pipeline models by 6.4% in PII detection accuracy.

Conference Publications

- Jiang, R., Daggett, J., Pingulkar, S., Zhao, Y., **Dhingra, P.**, Brown, D., Wu, Q., Zeng, X., Li, X., Xu, M., **BOE-ViT: Boosting Orientation Estimation with Equivariance in Self-Supervised 3D Subtomogram Alignment.** Paper Accepted at CVPR 2025. (Link)
- **Dhingra, P.** , Agrawal, S., Veerappan, C. S., Chng, E. S., Tong, R., **Enhancing Speech De-identification with LLM-Based Data Augmentation.** Paper Accepted at IEEE ICAICTA 2024. (Link)
- Veerappan, C. S., **Dhingra, P.** , Wang, D. Z., Tong, R., **SpeeDF - A Speech De-identification Framework.** Paper Accepted at IEEE TENCON 2024. (Link)
- **Dhingra, P.** ,Agrawal, S., Veerappan, C. S., Ho, T. N., Chng, E. S., Tong, R., **Speech de-identification data augmentation leveraging large language model.** Paper Accepted at IEEE IALP 2024. (Link)

Journal Publications

- **Dhingra, P.** , Agrawal, S., Veerappan, C. S., Chng, E. S., Tong, R., **Leveraging Large Language Models for Speech De-Identification.** Accepted at IJALP 2025. (Link)

Conference Presentations

- **IALP 2024:** International Conference on Asian Language Processing.
  Paper Presented: Speech de-identification data augmentation leveraging large language model
- **ICAICTA 2024:** International Conference on Advanced Informatics: Concepts, Theory, and Applications.
  Paper Presented: Enhancing Speech De-identification with LLM-Based Data Augmentation.

Teaching Experience

Rajiv Gandhi Institute of Petroleum Technology Jan 2024 - Apr 2024

Teaching Assistant: Computer Programming

- Conducted weekly lectures and hands-on coding sessions in C and Python for 45+ junior students.
- Guided students through 100+ algorithmic problems, fostering problem-solving skills and computational thinking.

Rajiv Gandhi Institute of Petroleum Technology Sept 2023 - Nov 2023

Teaching Assistant: Real Analysis and Calculus

- Delivered weekly lectures on Real Analysis and Calculus, breaking down complex mathematical concepts for 45+ junior students.
- Facilitated interactive problem-solving sessions, enhancing students' analytical skills and mathematical reasoning abilities.

Project Work

- **MultiLingual NER for Indian Languages using IndicBERT (2024):** Developed a Named Entity Recognition system for Indian languages using IndicBERT, achieving F1 score of 0.84. Implemented tokenization and label alignment techniques to process multilingual data, and utilized the ALBERT transformer architecture for pre-training language models on IndicCorp.
- **Sound Scape (2023):** Developed a spatial audio experience system integrating multi-channel processing and ML for environmental sound classification. Implemented audio feature extraction and stereo sound manipulation with distortion removal techniques.
- **Advanced 3D Animation: Bringing a Game Scene to Life with Maya (2022):** Developed a complex 3D animated game scene using Autodesk Maya, implementing character modeling, rigging, and environment creation. Utilized Python scripting for custom animations and implemented advanced rendering pipeline using Arnold renderer for high-quality output.

Skills

Languages: C/C++, Java, Python, JavaScript, SQL

Technologies & Tools: PyTorch, Git, GitHub, Docker, Kubernetes, Kafka, Jenkins, Flask, Pandas, Numpy,

Achievements

- Qualified GATE Computer Science and Information Technology (CS) 2024
- Qualified GATE Data Science and Artificial Intelligence (DA) 2024
- Qualified and attended Amazon Summer School 2023
- Semi Finalist at Flipkart GRiD 5.0 - Software Development Track (Top 1.5% of 400K)
- Achieved a top 2% rank in the IIT-JEE Advanced 2021 exam (1M+ candidates)