Shouvon Sarker

Curriculum Vitae

702 Santee Street, Apt. 7312
Prairie View, TX 77445
(337) 419-4543
shouvonsarker@gmail.com
shouvon-sarker
shovon095

Research Interests

Large Language Models (LLMs) · Neuro-Symbolic AI (Text-to-SQL) · Bayesian Deep Learning
Knowledge Distillation · Clinical NLP · Trustworthy AI (XAI) · Counterfactual Reasoning

Education

Ph.D. in Electrical Engineering, Prairie View A&M University, Texas, USA, *Jan 2023–Present*
Dissertation (Tentative): *Enhancing Structured Predictions in Large Language Models*
Advisors: Dr. Xishuang Dong, Dr. Lijun Qian

M.S. in Electrical Engineering, Prairie View A&M University, Texas, USA, *Aug 2021–Dec 2022*
Dissertation: *Medication Events Classification from Electronic Health Records Using BERT Models*
Advisors: Dr. Xishuang Dong, Dr. Lijun Qian

B.S. in Electronics and Communication Engineering, Khulna University of Engineering and Technology, Bangladesh, *Apr 2014–Mar 2018*

Selected Coursework

Selected Topics in Deep Learning: Bayesian Networks, Variable Elimination, GANs
Modern Artificial Intelligence: Constraint Satisfaction Problems, Optimal Decisions in Games, Hidden Markov Models

Publications

Manuscripts Under Review

[1] **Sarker, S.**, Qian, L., Dong, X. "From Tokens to Transitions: A Structured Jensen–Shannon Knowledge Distillation Method for NER." Submitted to *IEEE Transactions on Knowledge and Data Engineering*, 2025.

[2] **Sarker, S.**, Qian, L., Dong, X. "Learning SQL Correctness: Bayesian Error Localization, Counterfactual Repair, and Policy Optimization." *In Preparation*, 2025.

Conference Proceedings (Peer-Reviewed)

[3] **Sarker, S.**, et al. "Integrating Non-Parametric Attention to Enhance LLM-Based Text-to-SQL Without External Knowledge." *ICDM 2025*.

[4] **Sarker, S.**, et al. "Enhancing LLM Fine-tuning for Text-to-SQLs by SQL Quality Measurement." PhD Forum, *ICDM 2025*.

[5] **Sarker, S.**, Dong, X., Qian, L. "Text Generator and Text Discriminator for NIST GenAI T2T Challenge." *AIRC 2025*.

[6] **Sarker, S.**, Li, X., Dong, X. "Medical Data Augmentation via ChatGPT: A Case Study on Medication Identification and Medication Event Classification." *IEEE BHI 2023*.

[7] **Sarker, S.**, Dong, X., Qian, L. "Ensemble BERT for Medication Event Classification on Electronic Health Records." *ICIBM 2023*.

[8] Kuo, M., **Sarker, S.**, Qian, L., et al. "Enhancing Deep Knowledge Tracing via Diffusion Models for Personalized Adaptive Learning." *ASEE 2024*.

[9] Dong, X., **Sarker, S.**, Qian, L. "Integrating Human-in-the-loop into Swarm Learning for Decentralized Fake News Detection." *IDSTA 2022*.

Poster Presentations

[1] "Improving LLM-based Text-to-SQL Through Integrating Self-Discover Reasoning." *NASA DEAP Annual Meeting*, 2024.

[2] "Classification of Medication Events from Electronic Health Records Using BERT Models." *AIHC*, Rice University, 2024.

Research Experience

Graduate Research Assistant, Prairie View A&M University, *Jan 2023–Present*

- **Bayesian Text-to-SQL:** Hierarchical error localization with counterfactual repair and PPO optimization, improving execution accuracy by ∼12%.
- **Non-Parametric Attention:** Schema-grounded attention improving query accuracy by ∼10% without external knowledge.
- **Explainable Distillation:** Structured Jensen–Shannon divergence for interpretability in NER.
- **Deployment:** Interactive Text-to-SQL system for scientific databases.

Team Lead (NIST GenAI T2T Challenge), Prairie View A&M University, *2024*

- Designed a generative–critic framework for robustness evaluation.
- Achieved **Top-3 Global Placement** (Generator Track).
- Invited speaker, NIST GenAI T2T Workshop 2024.

AI Assistant Developer, AMIE 2025 Conference, *2025*

- Developed an Android AI assistant using OpenAI GPT-4o with a domain-specific knowledge base.

Graduate Research Assistant, Prairie View A&M University, *Aug 2021–Dec 2022*

- Ensemble BERT models for clinical NLP with calibration-aware prediction.

Teaching & Mentorship

Instructor, Prompt Engineering Workshop, *2024*

- Graduate-level instruction on Chain-of-Thought, ReAct, and applied LLM systems

Research Mentor, *2023–Present*

- Supervised undergraduate ROTC students

Awards & Honors

- Top 3 Placement (Top 10%), NIST GenAI Text-to-Text Challenge (2024)
- Outstanding Student Award, PVAMU CREDIT Center (2024)
- Top 10%, n2c2 Clinical NLP Challenge (2022)

Technical Skills

Languages: Python (Advanced), C/C++, SQL, HTML/CSS, PHP

Deep Learning: PyTorch, TensorFlow, Keras, Hugging Face Transformers

LLM Tools: LoRA, Adapters, PEFT, OpenAI API, LangChain

Data Science: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn

Databases: MySQL, PostgreSQL, SQLite

Deployment: Git, Docker, Linux, AWS

References available upon request.