Jillur Rahman Saurav

sauravsust71@gmail.com
linkedin.com/in/jillur-rahman-saurav — github.com/facevoid
facevoid.github.io — Google Scholar

Education

- **University of Texas at Arlington** — Arlington, Texas
  PhD in Computer Science — *August 2021 – Dec. 2025 (Expected)*
- **Shahjalal University of Science & Technology** — Sylhet, Bangladesh
  B.Sc. (Engg.) in Computer Science — *January 2013 – June 2018*

Experience

- **University of Texas at Arlington** — Arlington, TX, USA
  Research Assistant — *June 2022 - Present*
  - **Artificial Biomarker Stains**: Developed a cGAN architecture for generating missing biomarker channels in multiplexed spatial proteomics images, improving cost-effectiveness and efficiency in clinical diagnostics. Technologies: PyTorch, TensorFlow
  - **Histopathology Image Compression**: Engineered a VAE-based approach achieving a 1:512 compression ratio for cancer pathology slides while maintaining clinical accuracy, surpassing previous state-of-the-art. Technologies: PyTorch, scikit-learn
  - **Histopathology Slide Search Engines**: Conducted a comprehensive evaluation of four state-of-the-art histopathology slide search engines, assessing their clinical readiness and proposing requirements for successful clinical adoption. Technologies: Pytorch
- **University of Texas at Arlington** — Arlington, TX, USA
  Teaching Assistant — *August 2021 - May 2022*
  - **Artificial Intelligence**: Instructed and evaluated assignments for 120+ graduate students in CSE 5360. Topics: machine learning, neural networks, Agents modeling. Technologies: Python, TensorFlow, Keras
- **Pipilika** — Bangladesh
  Software Engineer — *Nov 2017 - Dec 2020*
  - **Natural Language Processing**: Developed deep learning models for Bangla language: speech recognition (3.96% WER), word prediction (trie-LSTM-N-gram hybrid), Part-of-Speech (POS) tagging (93.86% accuracy), and Named Entity Recognition (NER) (86% Macro-F1). Contributed to B-NER dataset (22,144 annotated sentences). Technologies: Python, TensorFlow, PyTorch, NLTK
  - **News and Data Processing**: Designed scalable news aggregator (10,000+ daily articles); built knowledge graph (5,552 websites); contributed to largest Bangla n-gram corpus and sentiment analysis dataset. Technologies: Django, Scrapy, Elasticsearch, Keras, Docker, Java, MongoDB
  - **Search Engine Optimization**: Implemented context-aware spell checker, query analysis with stemmer, and query expansion method. Improved overall search relevance and processing speed. Technologies: Spring Boot, Apache Solr, Keras, Python, Gensim
  - **Data Analytics and Research**: Conducted gender bias study in cross-cultural online newspapers; developed dengue surveillance system using news data; performed statistical analyses on COVID-19 self-screening tool (535,291 participants). Technologies: Python, Pandas, scikit-learn

Notable Publications

1. **Saurav, J.R.**, et al. (2023). A SSIM Guided cGAN Architecture For Clinically Driven Generative Image Synthesis of Multiplexed Spatial Proteomics Channels. *IEEE CIBCB*.
2. Nasr, M.S., ..., **Saurav, J.R.**, et al. (2024). Histopathology Slide Indexing and Search — Are We There Yet?. *NEJM AI*.
3. Nasr, M.S., ..., **Saurav, J.R.**, et al. (2023). Clinically Relevant Latent Space Embedding of Cancer Histopathology Slides through Variational Autoencoder Based Image Compression. *IEEE ISBI*.
4. Haque, M.Z., ..., **Saurav, J.R.**, et al. (2023). B-NER: A Novel Bangla Named Entity Recognition Dataset With Largest Entities and Its Baseline Evaluation. *IEEE Access*.

5. **Saurav, J.R.**, et al. (2021). A Comparative Study of Language Dependent Gender Bias in the Online Newspapers of Conservative, Semi-Conservative and Western Countries. *HCII*.
6. Tasnim, N., ..., **Saurav, J.R.**, et al. (2021). Observing the Unobserved: A Newspaper Based Dengue Surveillance System for the Low-Income Regions of Bangladesh. *FLAIRS-34*.
7. Sarker, S., ..., **Saurav, J.R.**, et al. (2020). Word Completion and Sequence Prediction in Bangla Language Using Trie and a Hybrid Approach of Sequential LSTM and N-gram. *ICAICT*.
8. Islam, M.R., **Saurav, J.R.**, et al. (2020). Query Expansion for Bangla Search Engine Pipilika. *IEEE TENSYMP*.
9. **Saurav, J.R.**, et al. (2019). End to End Parts of Speech Tagging and Named Entity Recognition in Bangla Language. *ICBSLP*.
10. **Saurav, J.R.**, et al. (2018). Bangla Speech Recognition for Voice Search. *ICBSLP*.

Skills

Languages:	Python, Java
Technologies:	PyTorch, TensorFlow, Keras, Elasticsearch, Django, Docker, Scrapy, Apache Solr, SQL, Spring Boot
Machine Learning:	Deep Learning, Computer Vision, GANs, VAEs, Natural Language Processing, Diffusion Models, Image Retrieval, Biomedical Image Analysis
Data Science:	Pandas, NumPy, scikit-learn, Data Visualization
Tools:	Git, Linux, Weights & Biases, LaTeX

Awards and Certification

- **Google Research CSRMP Recipient** — Class A
  Google Research — *January 2023 - April 2023*
- **Deep Learning Specialization** — Coursera
  Certification — *June 2019*