SHOIBOLINA KAUSHIK
GA, USA | +1 (510) 675-7443 | shoibolina.kaushik@gmail.com
LinkedIn| Website| ORCiD

SKILLS

- Languages: Python, SQL, R, C++, Java, Go
- ML/AI: PyTorch, Tensorflow, Transformers (Vision, NLP), Scikit-learn, NLTK, spaCy
- Other: Flask, Django, React.js, d3.js, PostgreSQL, AWS, GCP, Git

RESEARCH EXPERIENCE

Department of Environmental Sciences, Emory University, USA **Jul 2025 – Present**
Research Specialist

- Working in an interdisciplinary team on Gates Foundation funded project of malaria vector control in Ethiopia
- Automated processing and documentation of geospatial features by integrating 2500+ tabular data and imagery into structured format using Python libraries
- Prepared training datasets through annotation and data conversion in ArcGIS Pro, and supported baseline experiments for spatial detection models

Emory Center for Digital Scholarship (ECDS), Emory University, USA **Jan 2025 – May 2025**
Digital Scholarship Assistant

- Developed a novel pipeline in a 4-person, cross-functional team for enabling scalable analysis for 150+ years of historical maps of Atlanta, reduced manual digitization effort by over 95%
- Enhanced structural quality predictions by 41%, reducing geometric misalignments compared to random baseline
- Validated network consistency with 36% longer continuous road segments and 26% fewer fragmented components, demonstrating the pipeline's ability to preserve road topology for GIS workflows

Department of Biomedical Informatics, Emory School of Medicine, Emory University, USA **Jan 2024 – Jul 2024**
Graduate Research Assistant

- Collaborated in a 4-person team with physical therapists to design cost-effective, clinically applicable mobility assessment tool
- Generated human pose landmarks (keypoints) from 700+ smartphone videos (~250 GB) using Python based computer vision libraries and securely stored datasets on Emory's HPC cluster
- Delivered proof-of-concept classifier with 90% accuracy, later extended by team for multi-modal integration

Department of Computer Science, Manipal Institute of Technology, India **Aug 2021 – Dec 2023**
Undergraduate Research Assistant

- Developed a lightweight, integrable EEG classifier, that enabled rapid seizure detection with 97% accuracy within 3 minutes
- Collaborated with the Department of Pharmacy to build a diabetic treatment cost-prediction model with 96% accuracy
- Implemented blockchain based credit scoring algorithm in Go for accelerating agro-financing workflows in Southern India

PUBLICATIONS

- Unveiling the Epilepsy Enigma: An Agile and Optimal Machine Learning Approach for Detecting Inter-Ictal state from Electroencephalogram signals. Springer, 2024. DOI
- Classifying Simulated Gait Impairments using Privacy-preserving Explainable Artificial Intelligence and Mobile Phone Videos. Accepted, PLOS Digital Health. arXiv PDF
- Blockchain Framework for Agro Financing of Farmers in South India. Springer, 2023. DOI

HONORS & AWARDS

- Best Poster Award – Bio-STAR AI Symposium hosted by Wallace H. Coulter Department of Biomedical Engineering at Emory University and Georgia Institute of Technology, 2024.
- 50% tuition scholarship award, merit-based – Emory University, 2023-2025.

SELECT ACADEMIC PROJECTS

Global Burden of Diseases dashboard using D3.js (git demo) **May 2025**

- Designed an interactive D3.js choropleth dashboard visualizing global DALY rates across 200+ countries and 30+ disease categories using IHME datasets

- Provided spatiotemporal exploration of global health trends with dynamic country filtering and year-over-year disease comparisons

Biomedical Question Answering System (demo link) **Oct 2024 – Dec 2024**

- Built a RAG style PubMedBERT based biomedical QA system, fine-tuned on the BC5CDR dataset, achieving F1: 62% / Exact Match: 48% in extracting chemical-disease relationships
- Deployed interactive demo on HuggingFace Spaces with < 2s response time, enabling real-time query resolution

Synthetic iEEG Data Production for Epilepsy Analysis (git repo) **Oct 2023 – Dec 2023**

- Developed a Denoising Diffusion Probabilistic Model (DDPM) to generate synthetic 16-channel iEEG samples at 400 Hz
- Validated synthetic data fidelity through 5-fold cross-validation and spectrogram analysis (0-40 Hz), with classifiers trained on real data achieving AUC 0.73 (AdaBoost) and F1 0.54 (Naïve Bayes) on synthetic signals

EDUCATION

Laney Graduate School, Emory University, GA, USA **May 2025**
Master of Science in Computer Science **(GPA 3.72/4.0)**

Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), India **May 2023**
Bachelor of Technology in Computer Science and Engineering, *Minor Specialization in Big Data Analytics* **(CGPA 9.04/10)**

SERVICES

- Graduate Teaching Assistant, CS 377 Database Systems – Assisted Dr. Andreas Züfle, Fall 2023.
- Secretary, OpenBMI Outreach Club — Involved in organizing Atlanta Science Festival, Biomedical Career Panel, 2024-2025.
- Board Member, Rotaract Club – Directed fundraising initiatives; dedicated 100+ hours of community service, 2019-2023.
- Volunteer, TEDxMAHE -Countdown – spread awareness on climate change and interacted with the guests of honor, 2020.

MEDIA COVERAGE

- Featured for Best Poster Award at Bio-STAR AI Symposium 2024 — Emory University [post]