Ayaazuddin Mohammad

+1-857-423-9738 mohammad.ay@northeastern.edu https://linkedin.com/in/ayaazuddin https://github.com/ayaazuddin

EDUCATION

Northeastern University **Sep 2023 - May 2025**
Masters of Science, Artificial Intelligence

Vellore Institute of Technology **Jun 2019 - May 2023**
Bachelors of Technology, Computer Science and Engineering

TECHNICAL SKILLS

- **Programming Languages**: Python, Java, R, MATLAB
- **Web Development**: HTML, CSS, Javascript, SQL, React.js, Flask
- **ML/DL Frameworks**: PyTorch, TensorFlow, Keras, HuggingFace, scikit-learn
- **Development Tools**: Git, Linux, Docker, Docker Compose, PostgreSQL
- **Libraries & Tools**: OpenCV, NLTK, NumPy, pandas, OpenAI API

EXPERIENCE

Northeastern University **May 2024 - Present**
Graduate Research Assistant *Boston, MA*

- Designed a multimodal fair classifier for ED/ICU decision-making in **PyTorch** using models - **GPT-2** (HuggingFace Transformers) for object detection based report generation and **MedBERT**, improving clinical prediction fairness by 15%.
- Built a custom ensemble model entirely in **PyTorch** that improved demographic parity by 5% without compromising **F1-score**
- Quantified trade-offs in model fairness vs. performance across **3 modalities (images, text, structured data)**, identifying intersectional bias hotspots across 4+ demographic groups
- Submitted research findings to **IJCAI 2025**, demonstrating novel bias mitigation strategies in multimodal AI.

PUBLICATIONS

Mohammad, A., *et al.* (2025). *The Multimodal Paradox: How Added and Missing Modalities Shape Bias and Performance in Multimo dal AI*. Accepted at **CVPR 2025**, 2nd Workshop on Responsible and Generative AI (ReGenAI)

Mohammad, A., *et al.* (2024). *Fairness at Every Intersection: Uncovering and Mitigating Intersectional Biases in Multimodal Clinical Predictions*. arXiv preprint arXiv:2412.00606

PROJECTS

Water Quality Integration in ML-Based Prediction of Vector-Borne Diseases

- Led a study integrating water quality data into machine learning models for more accurate dengue prediction
- Proficiently harmonized different datasets from Indian government websites for **13 districts from 2015-2022.**
- Realized boost in performance with the integration of water quality data into the **Support Vector Machine** model surpassing the existing benchmarks in current research

AdaptAI: A Structured, Adaptive AI Communication Tool for Adults with Autism Spectrum Disorder

- Built full-stack web app **(Flask, React, PostgreSQL)** supporting ASD adults, containerized via Docker Compose for **100% reprod ucible environments**
- Integrated OpenAI's **GPT3.5-turbo** for adaptive content generation with custom prompting layer for ASD-friendly interactions
- Implemented **responsive UI** with HCI-compliant accessibility features and containerized the entire application stack using **Docker** for consistent development and production environments.

Unpacking Medical Texts: Analyzing Domain-Specific Behavior of Transformer Models on Reddit and PubMed

- Investigated how linguistic patterns (conversational vs. formal) impact medical text processing across Reddit and PubMed
- Benchmarked **BioBERT, ClinicalBERT, Reddit-tuned BERT, and BERT-base** from **Hugging-face** on 51k+ MedRedQA samples using classification and masked word prediction tasks.
- Achieved >97% F1 in source classification; applied **SHAP** to reveal source-specific linguistic attention and proposed domain adaptation strategies for healthcare NLP

Fairness at Every Intersection: Mitigating Biases in Multimodal Clinical Predictions

- Built unified text embeddings from 4+ EHR modalities using **MedBERT, ClinicalBERT, and BioClinicalBERT**
- Designed SDAE method, boosting fairness metrics (Demographic Parity ↑4–7%, TPR ↑3–5%) across 6+ race-gender subgroups
- Benchmarked on 2 datasets (MIMIC-Eye, MIMIC-IV ED) across 3 clinical tasks, maintaining 0.92+ F1 scores
- Outperformed Reject Option Classification by 10%+ in subgroup fairness without performance drop.
- Scaled to 51k+ multimodal patient samples, robust to missing modality settings.