Mahbuba Tasmin

PhD Candidate in Computer Science (Computational Biology & Machine Learning), UMass Amherst

413-479-9565 | mtasmin@umass.edu | linkedin.com/in/mahbuba-tasmin | github.com/Tasmin153 | Google Scholar

EDUCATION

University of Massachusetts Amherst — Amherst, MA
Ph.D. Candidate in Computer Science (Advisor: Prof. Anna Green) — *Expected Graduation: May 2027*
- Research Interests: Resistance Forecasting, Explainable ML, Protein Language Models, Structural Bioinformatics
- CGPA: 3.9/4.0
- Awards: Sudha and Rajesh Jha Scholarship, 2023
- Relevant Coursework: Advanced Algorithms, Machine Learning, Neural Networks, Computational Biology, AI, Information Assurance, Computer Architecture; Teaching for Tomorrow's Faculty; ML for Biological Sequence Data

University of Massachusetts Amherst — Amherst, MA
M.S. in Computer Science — *Sep. 2022 – May 2025*
- Thesis aligned with Ph.D. research on antibiotic resistance modeling.
- CGPA: 3.9/4.0

North South University — Dhaka, Bangladesh
B.S. in Computer Science and Engineering (Summa Cum Laude) — *Jan. 2016 – Dec. 2019*
- Concentration: Artificial Intelligence and Algorithms; CGPA: 3.89/4.0

RESEARCH EXPERIENCE

Graduate Research Assistant — Sep. 2022 – Present
SAGE Lab, University of Massachusetts Amherst — *Amherst, MA*
- Lead researcher on **BIG-TB**: a multimodal benchmark dataset (~17K isolates) for *M. tuberculosis* antibiotic resistance prediction across 11 WHO-priority drugs.
- Designed sequence- and structure-aware models: CNNs, Transformers, and fused-ridge baselines using DNA/protein features, and ESM embeddings.
- Constructed multi-species augmentation pipelines (UniProt, InterPro) to enhance protein generalization via evolutionary data.
- Performed explainability analyses with SHAP and causal variant recovery (recall@k) against WHO 2023 catalog.
- Lead researcher on **Forecasting Antibiotic Resistance Using Biophysics and Machine Learning**, supporting integrative modeling of resistance-conferring variants using protein thermostability and machine learning.
- Collaborated with cross-institutional researchers (Harvard DBMI, Farhat Lab) on multi-gene model reproducibility and benchmarking.
- Formulated and optimized a **Fused Ridge Regression** framework with convex regularization, fusion penalty, and enhanced gradient descent (momentum, clipping, Nesterov) for protein structure–aware resistance modeling.

INDUSTRY EXPERIENCE

AI Engineer — Mar. 2022 – Jul. 2022
NITEX Solutions Ltd. — *Dhaka, Bangladesh*
- Implemented Detectron2-based instance segmentation for product identification and OCR-based automation tools.
- Built fashion trend moodboards combining NLP and computer vision pipelines for workflow automation.

Software Engineer (AI & ML) — Jul. 2020 – Feb. 2022
M2SYS Technology — *Dhaka, Bangladesh*
- Developed image spoofing detection and contextual recommendation systems using ML and NLP techniques.
- Automated backend workflows with Camunda and deployed production ML models across distributed systems.

Selected Publications

Green, A. G., **Tasmin, M.**, Vargas, R., Farhat, M. R.
The structural context of mutations in proteins predicts their effect on antibiotic resistance.
Submitted to eLife. Preprint: bioRxiv 2025.09.23.676583 (2025)

Tasmin, M., Green, A.
Beyond Sequence-only Models: Leveraging Structural Constraints for Antibiotic Resistance Prediction in Sparse Genomic Datasets.
ICLR 2025 MLGenX Workshop (2025)

Yang, Z., Yao, Z., **Tasmin, M.** et al.
Unveiling GPT-4V's hidden challenges behind high accuracy on USMLE questions.
J Med Internet Res (2025)

Talks and Presentations

BIG-TB: A Benchmark Dataset for Genomic Resistance Prediction and Interpretability in *Mycobacterium tuberculosis*
Machine Learning for Computational Biology (MLCB) Workshop, 2025
Spotlight talk presented highlighting dataset design and explainability analyses.

Protein Structure-Informed Regularized Linear Model Outperforms ESM for Predicting Antibiotic Resistance
Program in Quantitative Genomics Conference (PQG), Harvard University, 2024
Poster presentation demonstrating integration of 3D structural features with machine learning models.

Teaching Experience

Head Teaching Assistant — Fall 2025 – Spring 2025
CompSci 520: Theory and Implementation of Advanced Software Engineering — *UMass Amherst*
- Led 140+ students, managed teaching assistants, and coordinated course logistics.
- Maintained GitHub Classroom and Gradescope automation pipelines.

Course Developer Assistant — Summer 2023
CompSci 520: Theory and Implementation of Advanced Software Engineering — *UMass Amherst*
- Revamped course structure, labs, and assignments with automated grading.

Honors and Awards

CRA-WP Grad Cohort for Women — 2023
San Francisco, USA

Technical Skills

Tools & Languages: Python, Bash, R, LaTeX, pandas, NumPy, matplotlib, Git, Docker, Linux, SLURM
ML/AI: PyTorch, scikit-learn, XGBoost, CNNs, Transformers, Random Forests, SHAP
Bioinformatics: UniProt, InterPro, Rosetta, AAIndex, Protein Structure, VEP/ANNOVAR

Leadership & Mentorship

PhD Graduate Representative — Fall 2025
Student Representative in Faculty Senate — *College of Computer & Information Sciences, UMass Amherst*

Shakir Sahibul — Fall 2024
M.S. Student, UMass Amherst — *Supervised Transformer-based resistance prediction project.*

Suqi Hong — Fall 2025
M.S. Student, UMass Amherst — *Supervising EvoAug-based protein resistance prediction project.*