

# Sowmya Muthyala

Boston, MA | [muthyala.s@northeastern.edu](mailto:muthyala.s@northeastern.edu) | 857-370-1821 | [LinkedIn](#) | [Github](#) | [Portfolio](#)

## EDUCATION

**Northeastern University**, Boston, MA

**Dec 2024**

Master of Science in Bioinformatics

**Sreenidhi Institute of Science and Technology**, Hyderabad, India

**Aug 2020**

Bachelor of Technology in Biotechnology

## SKILLS

**Programming & Scripting:** Python (Biopython, NumPy, Pandas, Matplotlib, scikit-learn, PyTorch), R (Bioconductor, dplyr, ggplot2), Shell, BASH, SQL, Perl, MATLAB

**Bioinformatics Tools & Version control:** STAR, DESeq2, Salmon, rMATS, Trinity, BLAST, FASTQC, Samtools, Bowtie2, MEME Suite, PRIESSTESS, FIMO, SPAdes, Picard, bcftools, BWA, IGV, GSNAP, JIRA,

**Workflow Management & Cloud Platforms:** Docker, Conda/Mamba, SLURM, Git/GitHub, UNIX/Linux, SSH, AWS, Azure

**Data Science & Data Visualization:** Statistics, Model Evaluation, Tableau, Power BI, algorithm development, process optimization

**Domain Knowledge:** RNA-seq, eCLIP, Differential Expression, Motif Discovery, Transcriptome Assembly, Functional Genomics, Next-Generation Sequencing (NGS), Protein Expression & Purification, troubleshooting, database management

## WORK EXPERIENCE

**Bioinformatician and Machine Learning** - Institute for Experiential AI, Boston

**Jul 2024 - Present**

- Developed high-throughput RNA-seq analysis pipelines using STAR and rMATS, enabling the processing of over 500 samples on HPC clusters and improving pipeline efficiency by 30%
- Identified RNA-binding motifs from eCLIP data by applying the PRIESSTESS framework and fine-tuning model parameters, enhancing motif resolution and reproducibility across datasets
- Performed differential gene expression analysis using Salmon and DESeq2, revealing 200+ differentially expressed genes relevant to immune activation pathways (FDR < 0.05)
- Standardized and documented analysis workflows, facilitating knowledge transfer and improving reproducibility across multi-disciplinary teams

**Quality Control Microbiologist** - Revelations Biotech Pvt Ltd, India

**Jan 2021 - May 2022**

- Streamlined microbial quality testing workflows, automating over 300 tests/hour with robotic systems and improving lab throughput by 40%
- Enhanced enzyme production processes through data-driven quality control, boosting fermentation efficiency by 4% while maintaining regulatory compliance
- Analyzed over 740 samples weekly using statistical and microbiological techniques, ensuring consistency, traceability, and adherence to quality standards
- Authored detailed documentation protocols that supported GMP readiness and improved cross-functional reporting

**Project Intern** - Center for Biotechnology JNTU, India

**Jan 2020 - Apr 2020**

- Conducted phytochemical profiling of commercial fruit juices, achieving 94% accuracy across 140+ samples using UV-Vis spectrophotometry
- Designed reproducible experimental protocols to evaluate six antioxidant compounds, contributing to research on nutritional quality in food biotechnology

**Summer Intern** - Indian Institute of Technology, Hyderabad, India

**Jun 2019 - Aug 2019**

- Engineered recombinant protein expression systems by cloning bacterial genes, improving production yield in targeted constructs
- Performed over 20 rounds of protein purification using ion exchange, size exclusion, and affinity chromatography, significantly increasing average yield and consistency

## ACADEMIC PROJECTS

**Heart Disease Prediction Using Machine Learning Models**, Northeastern University

**Dec 2023**

- Built predictive models using SVM, Random Forest, and Logistic Regression to classify disease risk from clinical datasets
- Achieved 73% prediction accuracy with SVM using cross-validation and confusion matrix evaluation
- Engineered features and conducted model tuning, showcasing practical application of ML in bioinformatics diagnostics

**De novo transcriptome assembly and analysis of bacteria**, Northeastern University

**Dec 2023**

- Designed a complete RNA-Seq pipeline from quality control to annotation, achieving high-confidence transcriptome assembly using Trinity on raw Illumina reads
- Identified and quantified differentially expressed genes using DESeq2, and annotated protein-coding sequences through BLASTp alignment against Swiss-Prot, enhancing downstream pathway analysis accuracy

## LEADERSHIP

**Teaching assistant** - Graduate Leadership Institute

- Mentored more than 50 students to refine their leadership approach and boost confidence in professional communication