

# Ayemen Maqsood Mulbagal

Boston, MA | +1 (857) 296-2150 | [ayemenmaqsood.2000@gmail.com](mailto:ayemenmaqsood.2000@gmail.com) | [Linkedin](#) | [github.com/mulbagalamaq](https://github.com/mulbagalamaq)

## EDUCATION

### Master of Science in Bioinformatics

December 2024

Northeastern University, Boston, MA

**Coursework:** Bioinformatics Computational methods, Statistics for Bioinformatics, Machine Learning and Data Mining, Bioinformatics Programming, Genomics in Bioinformatics

### Bachelor of Engineering in Biotechnology

August 2018 – August 2022

Ramaiah Institute of Technology, Bangalore, India

**Coursework:** Cell and Molecular Biology, Immunology, Structural Biology, Genomics, Proteomics, Bioinformatics, Biostatistics and Biomodelling and Biochemistry

## WORK EXPERIENCE

### Quantitative Biology Co-op

July 2024 – December 2024

Vor Bio, Boston, MA

- Analyzed public digenome-seq data to validate the in-house analysis tool kit for off-target nomination, uncovering six previously missed off-target sites that were overlooked by other public tool
- Conducted comparative analysis and quality control of digenome-seq data from in-house and public datasets, using IGV and UCSC genome browser to map over 260 cut sites, highlighting in-house tool's superior accuracy and reducing off-target edit risk
- Built scalable bioinformatics pipelines in Python, R, Bash, and AWS for high-throughput CRISPR Digenome-seq data processing, automating digenome and downstream analyses with Nextflow and custom scripts, cutting analysis time by 35% and reducing manual work through integrated logging, benchmarking, and email alerts
- Collaborated with Process Development (PD) and Cancer Immunotherapy (CIT) teams to participate in NovaSeq and MiSeq sample preparation, increasing accuracy in identifying NGS artifacts
- Optimized data visualization strategies with pandas, matplotlib, and seaborn to present genomic data trends effectively to cross-functional teams, achieving a 30% increase in clarity and interpretability of complex data

### RBP binding site analysis based on eClip Data

January 2024 – June 2024

Research Associate EAI, Northeastern University

- Built a preprocessing pipeline for eCLIP data incorporating demultiplexing, adapter trimming, sequence alignment, and PCR duplicate removal. Successfully identified RNA-binding protein peaks using the CLIPPER algorithm under the guidance of Dr. Ayan Paul.

### Cancer Systems Biology LAB, BSSE, Indian Institute of Science

February 2022 – June 2022

Research Intern, Bangalore, India

- Performed RNA-Seq analysis using Bioconductor in R and Python on a linux HPC cluster, processing over 150 cancer cell line samples to explore EMT dynamics and identified key gene targets through GSEA and constructed a regulatory network involving miR-200, SNAIL, ZEB-1, and SLUG, advancing insights into EMT-related gene regulation in oncology

## ACADEMIC PROJECTS

### Advanced Analytical Approach in Glioma Grading Using Clinical and Genetic Data

[Northeastern University]

#### Glioma Classifier GitHub

- Developed machine learning models (KNN, Naive Bayes, Logistic Regression, SVM) for glioma grading, achieving a 40% improvement in accuracy by integrating clinical and genetic data
- Processed and engineered features from large datasets and performed statistical validation (Chi-Square, ANOVA, cross-validation) in R. Built an ensemble model that outperformed traditional methods, increasing predictive accuracy for glioma grades by 25%

## TECHNICAL SKILLS

**Languages and Frameworks:** Python, R, SQL, Shell scripting, Biopython, Bioconductor, Sklearn, NumPy, Pandas, Seaborn, PyTorch

**Bioinformatics Skills:** Next Generation Sequencing (NGS) analysis, Genomics, Version Control (Git), Bioinformatics Databases, Biostatistics, Whole Genome Sequencing, Variant Calling, RNA-seq, CHIP-seq, Digenome-seq, Single-Cell Analysis (Scanpy)

**Bioinformatics tools:** BLAST, GSNAP, STAR, GATK, Trimmomatic, DESeq2, Trinity, QIIME, IGV, Bowtie, BEAST2, CLIPPER, Bedtools, FastQC, Cutadapt, UMI-tools, SAMtools, UCSC Genome Browser, BWA, Picard

**Tools and Technologies:** HPC, SLURM workload manager, CWL, Snakemake, Nextflow, AWS EC2, S3, Docker, Kubernetes, Plotly

**Laboratory Techniques:** Aseptic technique, Isolation and Identification of the microorganisms, Protein Assays, PCR, HPLC, blotting techniques, SDS page, Chromatography, GC-MS, ICP-MS

## LEADERSHIP

### Head of Bioinformatics, Graduate Bioinformatics and Biotech Association

August 2024 - Present

- Organized a bioinformatics hackathon with 40 participants, tackling industry-relevant challenges and leading workshops on command-line techniques and version control
- Launched a monthly journal club, fostering peer-led discussions on current bioinformatics research and methods
- Spearheaded partnerships with industry professionals, arranging guest lectures to bridge the gap between academic learning and real-world applications