ANAY ATHAWALE

Irving, TX | (617) 602-2831 | athawale.a@northeastern.edu | https://www.linkedin.com/in/anay-athawale-146586162/

PROFESSIONAL SUMMARY

Innovative and results-driven Data Scientist with a strong foundation in computational biology, bioinformatics and public health. Proficient in developing and deploying bioinformatics pipelines and tools, with a focus on machine learning and cloud computing. Passionate about advancing biomedical research through data-driven insights.

EDUCATION

Northeastern University, Boston, MA, USA   Grad: **12/2023**
Master of Science in Bioinformatics
Massachusetts College of Pharmacy and Health Sciences, Boston MA, USA   Grad: **05/2021**
Master of Public Health
Maharashtra University of Health Sciences, Aurangabad, Maharashtra, India   Grad: **03/2018**
Bachelor of Dental Surgery

TECHNICAL SKILLS

Programming Languages: Python, R, Bash Shell, Java, C++
Bioinformatics Tools: BLAST, GSNAP, Trimmomatic, SAMtools, BioPython, Clustal Omega
Cloud Computing Platforms: AWS
Pipelining Tools: Snakemake, Nextflow, Airflow
Containerization: Docker
Database Programming: SQL
Data Analysis Tools: Numpy, Pandas, Jupyter Notebook, BioPython
Cheminformatics Libraries: RDKit, OEchem
AI & Machine Learning Libraries: TensorFlow, Keras, PyTorch, Scikit-Learn, DeepChem
Databases: SQLite, SwissProt, NCBI, KEGG, GO, GenBank, PDB, Ensembl, chEMBL
Web/Front-end Development: React.js
Operating Environments: Linux, Windows, Mac

WORK EXPERIENCE

Moderna Therapeutics, Cambridge MA, USA   **07/2023 – 12/2023**
Bioinformatics Co-op
• Spearheaded the development of a novel pipeline for antibody sequence discovery, enhancing targeted therapy research by 30%.
• Engineered and managed genomic data databases, achieving a 25% improvement in data retrieval efficiency.
• Implemented cloud-based bioinformatics pipelines, optimizing NGS data processing by 20%.
• Conducted extensive analyses on antibody sequences, contributing significantly to the understanding of immune responses.
• Utilized machine learning models to predict antibody efficacy, achieving an 85% accuracy rate in preliminary tests.

Northeastern University, Boston MA, USA   **07/2022 – 07/2023**
Graduate Research Assistant - Cheminformatics
• Contributed to SARS-CoV-2 drug discovery research by assisting in the identification and analysis of potential drug compounds, enhancing the project's progress by 20%.
• Automated the processing of protein structures using YASARA and POOL for predicting active amino acid residues, increasing workflow efficiency by 30%.
• Developed an automated system for creating receptor grids around active sites and preparing structures for docking procedures using Schrödinger Glide, reducing processing time by 25%.
• Implemented and refined local docking procedures, improving accuracy in ligand-binding predictions by 15%.

Rutgers University, Newark NJ, USA   **07/2021 – 12/2021**
Clinical Preceptorship in Diagnostic Sciences
• Analyzed patient genomic data using bioinformatics techniques, improving diagnostic accuracy by 20%.
• Utilized libraries like Pandas, NumPy and Matplotlib for statistical analysis of dental imaging data, aiding in oral disease research.
• Conducted molecular marker analysis using BLAST and BioPython, identifying key genetic variations associated with oral diseases.
• Streamlined clinical data management and analysis, enhancing efficiency in diagnostics by 15%.

PROJECTS

Computational Drug Discovery,
• Developed a QSAR model for Acetylcholinesterase inhibitors, achieving an 80% accuracy in pIC50 value prediction.
Deployment of ML Model as Web App
• Developed a user-friendly web application using React.js for the front-end to predict the biological activity of compounds.
Genomic Data Visualization Tool
• Developed a Python-based tool for visualizing genomic data, improving data interpretability for research teams.
Automated Sequence Analysis Pipeline
• Designed an automated pipeline for DNA sequence analysis, reducing processing time by 25% and enhancing accuracy.