Anay Athawale

Irving, TX | (617) 602-2831 | athawale.a@northeastern.edu | LinkedIn | Github | Portfolio

SKILLS

Programming Languages: Python, R, SQL, HTML, CSS
Tools & Methodologies: Git, Agile, CI/CD, Jira, AWS, Docker, Azure
Libraries: TensorFlow, Keras, PyTorch, Scikit-Learn, Numpy, Pandas, Matplotlib, MATLAB
Domains: Machine Learning, Deep Learning, Artificial Intelligence, Data Science, Natural Language Processing, Computer Vision

WORK EXPERIENCE

Moderna Therapeutics, Cambridge, MA, USA **07/2023 – 12/2023**
Data Science Co-op

- Developed k-means clustering and Principal Component Analysis (PCA) models for antibody sequence discovery, enhancing targeted therapy research efficiency by 30% using Python, TensorFlow, and AWS.
- Optimized large genomic databases, achieving a 25% improvement in data retrieval times, leveraging SQL and data management skills on cloud environments.
- Implemented scalable cloud-based bioinformatics pipelines, enhancing next-generation sequencing data processing efficiency by 20%, using AWS, Docker, and Nextflow.
- Conducted statistical analyses and applied machine learning techniques, such as PCA and Scikit-Learn, on antibody sequences to advance the understanding of immune responses.
- Designed regression and Bayesian models for predicting antibody efficacy, achieving an 85% accuracy rate, using TensorFlow and Keras.

Northeastern University, Boston, MA, USA **07/2022 – 07/2023**
Research Assistant – DATA Lab

- Developed a multilingual AI chatbot for Northeastern University's Office of Global Services (OGS) using OpenAI API, Python, Beautiful Soup, and LangChain, improving information accessibility for international students.
- Created and deployed an NLP model for query visualization, enhancing data understanding and accessibility by employing techniques such as Named Entity Recognition (NER) and sentiment analysis using libraries like SpaCy and NLTK.
- Conducted research on data provenance and data lake management, implementing scalable data management solutions using Apache Spark and Hadoop to optimize data integration and discovery processes.
- Collaborated on projects involving scalable inference over uncertain and networked data, applying advanced statistical methods such as Bayesian inference and machine learning techniques including Random Forests and Gradient Boosting to solve complex data problems.

Rutgers University, Newark, NJ, USA **07/2021 – 12/2021**
Data Analyst – Clinical Preceptorship in Diagnostic Sciences

- Integrated clinical knowledge with genomic data analysis, utilizing machine learning models such as logistic regression and support vector machines (SVM) to improve diagnostic accuracies in oncology and refine treatment protocols.
- Streamlined data management processes by developing automated data pipelines using Python and SQL, enhancing clinical research efficacy and maintaining detailed patient records with increased accuracy and efficiency.

PROJECTS

Action Detection Using LSTM Neural Networks Github
• Implemented real-time action recognition using TensorFlow, OpenCV, and MediaPipe, training an LSTM model to interpret pose estimations and classify human actions from video input, providing real-time visual feedback.

Customer Survival Analysis and Churn Prediction App Github
• Conducted survival analysis with Kaplan-Meier curves and Cox-proportional hazard models to identify factors influencing churn, and developed a Random Forest prediction model within a Flask app, complete with visual analytics for strategic retention efforts.

Computational Drug Discovery and Deployment of ML Model as a Web Application Github

- Developed a QSAR model for Acetylcholinesterase inhibitors using advanced algorithmic techniques, such as Random Forest and Gradient Boosting, achieving 80% accuracy in pIC50 value predictions, thereby informing potential treatment targets.

EDUCATION

Northeastern University, Boston, MA, USA Grad: **12/2023**
Master of Science in Bioinformatics
Relevant coursework: Machine Learning, Data Mining, Collecting Sorting Retrieving Data

Massachusetts College of Pharmacy and Health Sciences, Boston MA, USA Grad: **05/2021**
Master of Public Health
Relevant coursework: Statistics, SAS, Epidemiology,

Maharashtra University of Health Sciences, Aurangabad, Maharashtra, India Grad: **03/2018**
Bachelor of Dental Surgery