Raghu Ram Sattanapalle

 Boston, MA, USA - 02119   +1 (347) 873-2177   sattanapalle.r@northeastern.edu
 linkedin.com/raghuramsattanapalle   github.com/RaghuRamSatt   raghuramsatt.github.io

Education

Northeastern University — Expected May 2025
Master's of Science in Data Science (GPA: 4.00 / 4.00) — *Boston, MA, USA*
- **Relevant Coursework:** Large-Scale Parallel Data Processing, Supervised Machine Learning, Database Management

New York University — May 2018
Master's of Science in Mechanical Engineering — *New York, NY, USA*
- **Relevant Coursework:** Robot Perception, Simulation Tools and Software for Mechatronics and Robotics

Technical Skills

Programming Languages: Python, R, MATLAB, C++, SQL, Scala, Julia, JavaScript
Machine Learning: TensorFlow, PyTorch, Keras, FastAI, Scikit-learn, Deep Learning, Time Series Analysis,
Data & Cloud Technologies: AWS, Hadoop, Spark, Hive, HBase, PostgreSQL, MySQL, MongoDB
Data Visualization: Tableau, Plotly, Matplotlib, Seaborn, Bokeh, ggplot2, Excel
Software Development: Git, Docker, CI/CD

Experience

Northeastern University, Khoury College of Computer Sciences — July 2024 - Present; Sept 2023 - Dec 2023
Graduate Teaching Assistant - CS5800 Algorithms — *Boston, MA*
- Facilitated graduate-level algorithms education for over 35 students every semester, covering topics from fundamental data structures to advanced algorithms, including Dynamic Programming, Graph Algorithms, and NP-complete problems.
- Provided comprehensive support through grading assignments, offering constructive feedback, and conducting weekly office hours, enhancing student understanding and performance.

Veeco Instruments — Jan 2024 - June 2024
Engineering Data Scientist (Co-op) — *San Jose, CA*
- Led the development of CNNs using TensorFlow and PyTorch to predict boron wafer resistance (RS), initially facing high average error rates of 13-22%.
- Reduced the RS prediction average error rate to 6% by implementing domain-specific data augmentation and group normalization techniques.
- Created Python scripts to automate the extraction and structuring of tool and wafer data from storage drives, achieving 100% data capture and generating comprehensive datasets for model training and analysis.
- Led the development of visualization tools for manufacturing tool channel data, enabling quick identification of trends and anomalies, which enhanced process control and informed decision-making.

NYU Dynamical Systems Laboratory — Jan 2022 - Aug 2022; June 2019 - Aug 2021; Sep 2018 - Jan 2019
Researcher/ Research Assistant — *Brooklyn, NY*
- Spearheaded the development of machine learning models to predict ICU patient mortality rates, achieving 90% accuracy and contributing to improved patient care strategies in healthcare analytics.
- Conducted groundbreaking causal analysis study on gun prevalence and mass shootings, resulting in a publication in Nature Human Behaviour and securing $2.1 million in research funding.
- Created innovative information theory-based models of zebrafish behavior, leading to a publication in Flow: Applications of Fluid Mechanics and advancing understanding of collective animal behavior.

Projects

Optiver - Trading at the Close: Predict US Stock Movements | *Python, LightGBM, XGBoost, PyTorch*
- Developed machine learning models to predict stock price movements during the closing auction for NASDAQ-listed stocks, achieving top 20% ranking in a Kaggle competition.
- Preprocessed over 5 million data points, applying feature scaling and data imputation to enhance model performance.
- Engineered advanced features including temporal lagged variables and rolling window statistics, significantly improving model accuracy in predicting stock price movements.
- Implemented and fine-tuned various models including LightGBM, XGBoost, and Neural Networks, optimizing for Mean Absolute Error (MAE).

Soundit: Database-Driven Music Streaming Platform | *MySQL, JavaScript, Python*
- Co-developed a dynamic music platform integrating complex MySQL database schema for user interactions and music management.
- Engineered a recommendation system for personalized user experiences, enhancing user engagement and platform functionality.
- Implemented user authentication, subscription services, and interactive features using JavaScript and Python.

Environmental Impact Analysis: Air Quality During COVID-19 | *Python, SQL, Data Visualization*
- Led a team investigation into the correlation between COVID-19 lockdown measures and air quality in Massachusetts, utilizing Air Quality Index (AQI) and Social Distancing Index (SDI) as key indicators.
- Leveraged Python and SQL to process and analyze data from multiple sources, creating insightful visualizations to communicate findings.
- Developed a comprehensive analysis framework that could be applied to similar environmental impact studies in urban areas.