Saunak Roy

Germantown, MD | 530-798-0597 | saunakr769@gmail.com | linkedin.com/in/saunak-roy | github.com/saunakroy

Education

University of Maryland, College Park — College Park, MD

BS, Computer Science: Machine Learning and Mathematics: Statistics — *Dec 2026*

Coursework: Multimodal Deep Learning, Data Science, Machine Learning, Computer Vision, Artificial Intelligence, Advanced Data Structures, Algorithms, Probability Theory, Computer Systems, Discrete Math, Differential Equations

Activities: University Honors, CS Department Honors, BigThinkAI (Data Science Lead), UMD AI Alignment Group

Experiences

AstraZeneca — Frederick, MD

Data Science Intern — *June 2025 - Present*

- Partnering with Manufacturing Sciences and Technology to perform modeling and visualization of process performance of commercial products, microbial monitoring, raw material data, etc.
- Applying statistical tools, **machine learning** techniques (anomaly detection) to analyze biologics manufacturing
- Built a Python data pipeline to streamline raw material data processing, contributing to **$12,670** in soft savings

OUR Immersive Research Internship Experience — College Park, MD

Research Intern — *May 2024 – Feb 2025*

- Used R and data science and analytics techniques to examine undergraduate AI literacy at UMD
- Discovered positive correlation (ρ=0.6) between students' technical AI knowledge and perception of AI's risks
- Paper accepted to **ACM SIGCSE 2025** conference

National Institute of Standards and Technology (NIST) — Gaithersburg, MD

Software Engineer Intern - Applied and Computational Mathematics Division — *Jun 2022 – Aug 2022*

- Studied non/minimal intersecting n-ary mathematical trees generated on smooth 2D manifolds (spheres, tori, etc.)
- Generated random, deterministic, symmetric trees, and trees with circular boundaries using **Java**
- Implemented Bentley-Ottmann sweep line algorithm and other scalable computational geometry algorithms
- Received **second-best research presentation** award out of **41** interns in August presentation session

Projects

Brain Tumor Classification Using CNNs | *Tensorflow, Keras, NumPy, Pandas, OpenCV, Scikit-learn, Matplotlib*

- Built a **deep learning** model that classifies brain tumor scans from 150x150 MRI images using a custom **CNN** and **transfer learning** through ResNet-50 and VGG models
- Incorporated data augmentation and **regularization** techniques to improve model accuracy, achieving **0.96** F1-score for CNN performance

Predictive Analysis of Trends in US Wealth Disparities | *NumPy, Pandas, Matplotlib, Scikit-learn, SciPy*

- Designed a tutorial for analyzing **49 years** (1973-2022) of hourly compensation data based on sociodemographics (race, gender, etc.) using **statistical analysis/machine learning** techniques
- Utilized regression, ANOVA, and t-tests to assess wage impacts and predict future trends, achieving R-squared values as high as **0.95**

Skills

Languages: Python, Java, C/C++, SQL (Oracle), OCaml, HTML/JS/CSS, R, XML, MATLAB
Libraries: NumPy, Pandas, Matplotlib, PyTorch, TensorFlow, OpenCV, Keras, Scikit-learn, Scipy, ggplot
Frameworks/Other tools: Next.js, Flask, AWS, FastAPI, Swing, Kali Linux
Awards: National Cyber Scholarship Competition Finalist (2021), USA Computing Olympiad Silver Division (2021), University of Maryland High School Math Competition Part II Qualifier (2021, 2022)