Raunak Srivastava

raunaksrivastava001@gmail.com | (408) 816-5744 | [LinkedIn](#) | [Portfolio](#)

Skills

- **Programming Languages:** Python, R, SQL, Java, SAS, MATLAB, C, Bash (Unix Shell)
- **Tools/Packages:** Pandas, NumPy, Hadoop, Spark, Scikit-learn, SciPy, Matplotlib, Plotly, Dash, Seaborn, Plotnine, Tableau, Excel, Git
- **Technical Skills:** Machine Learning, Artificial Intelligence, ETL, Data Visualization, Predictive Analytics, Feature Engineering, Model Evaluation and Validation, Agile Methodologies

Education

UNIVERSITY OF CALIFORNIA, IRVINE | SEP 2023 – DEC 2024

- **Master of Data Science**
- Coursework: Artificial Intelligence, Machine Learning and Data Mining, Probability and Statistical Theory, Databases and Data Management, Statistical Methods

PURDUE UNIVERSITY | WEST LAFAYETTE, IN | 2019-2023

- **B.S. in Data Science, Minor in Computer Science**, **Certificate in Entrepreneurship and Innovation**
- Relevant Coursework: Data Structures and Algorithms, Object Oriented Programming, Data Mining and Machine Learning, Artificial Intelligence, Large Scale Data Analytics, Computer Architecture, Numerical Methods, Statistics for Data Science, Linear Algebra, Probability, Statistical Programming

Experience

- **The National Institutes of Health | Data Science Intern** (May 2021 – May 2023)
  - **Dashboard Design and Implementation** – Developed 10+ interactive dashboards to provide real-time insights into key metrics ensuring user-centric interfaces using Plotly Dash and Python.
  - **NOAA Weather** – Extracted and preprocessed weather forecast data from the NOAA website to predict future campus load using the weather.gov API, PI System Explorer, and Python. Calculated and visualized error metrics (mean, standard deviation, MAD, and RMSE) of weather data using Plotly and prepared total performance overview dashboards to be integrated into internal NIH forecaster quality web app.
  - **Performance Testing CUP Models** – Testing, automation, statistical analysis, and visualization of the NIH Central Utility Plant (CUP) machine learning models using MATLAB, PI System Explorer, and Excel.
- **JPMorgan Chase & Co. | AI and Data Science Summer Analyst** (Jun 2022 – Aug 2022)
  - **Statistical Testing Automation** – Utilized Python and SQL to automate process of selecting stratified samples, performing McNemar Testing, summarizing, and visualizing results for HMDA Geocode Sourcing.
- **Onpoint Insights | Data Science Intern** (Jun 2020 – Aug 2020)
  - **Credit Card Fraud Detection** – Implemented and compared performance of 3 ML algorithms in Python (Random Forest, Neural Networks, and LightGBM) to identify and prevent credit card fraud transactions.
  - **E-Commerce Sentiment Analysis** – Used Multinomial Naïve Bayes Classifier to perform sentiment analysis in Python and designed dashboard in Tableau to display insights for client company.
- **The Data Mine** – **Purdue University | Undergraduate Researcher**
  - Teaching Assistant with corporate partner IU Health
  - Data visualizations for business development & post-COVID stadium seating plans with Purdue Athletics

Awards

- Corporate Partners Program Scholarship
- Dean's List and Semester Honors

Other

- US Citizen
- Point Guard of Purdue Men's Club Basketball Team