Shreya Banik

+1 (203) 676-7828 | sbani1@unh.newhaven.edu | shreya-insights | shreya.github | shreya.portfolio | New Haven, CT

Experience

University of New Haven — Connecticut, USA
Research Assistant — *Aug 2023 – Present, Part-time*

- Advance scientific research on clinical data by orchestrating *Microbiome Data Analysis* using Python & R-programming.
- Designed and deployed strategic data validation procedures while maintaining data consistency to improve data quality.
- Conducted an efficient data management process utilizing secure cloud storage to ensure compliance with data privacy policies.
- Collaborated on a research paper based on findings & documented insights to contribute to scientific literature in *Neuroscience*.
- Compared various machine learning models to identify the most effective approach for data analysis, optimizing performance and accuracy in clinical research applications.

Cognizant Technology Solution — Kolkata, India
Programmer Analyst — *Oct 2021 – Jul 2023, Full-time*

- Implemented a comprehensive *Automatic Reconciliation Tool* using MySQL Workbench and MS Excel to efficiently track over 400,000 daily transactions from ServiceNow to the database, resulting in an impressive 78% reduction in resource efforts.
- Documented policy and claims-related large datasets, providing actionable insights to technical and non-technical stakeholders.
- Applied SQL queries for data manipulation to uncover key-patterns, resulting in a 30% improvement in the reporting accuracy.
- Identified and reduced defects by 8X, showcasing effective problem-solving skills and process improvement.

CBNITS Private Limited — Kolkata, India
Data Analyst — *April 2021 – September 2021, Intern*

- Performed end-to-end research to inform AI-driven decision-making processes, improving overall project outcomes and introducing new app development classes in finance that enhanced the application flexibility and functionality by 68%.
- Analyzed large financial datasets on cryptocurrency exchange to derive key insights and integrated data pipeline using Rest APIs, MySQL and Microsoft OneDrive for cloud-based data storage.
- Teamed up with cross-functional teams to resolve critical problems in the application by maintaining Excel reports and adding innovative formulas to boost the speed of operation by 20%.

Education

University Of New Haven — West Haven, CT
MS in Data Science — *Aug 2023 – May 2025*

- Relevant Coursework: Data Science, Artificial Intelligence, Machine Learning, Distributed and Scalable Data Engineering, Computer Vision, Deep Learning, Natural Language Processing, Bayesian Data Analysis
- Dean's Scholarship, *Industrial Project Poster Presentation*, Volunteer Capstone Project Completion

Hooghly Engineering & Technology College — Kolkata, India
B.Tech in Electronics and Communication and Engineering — *Aug 2017 – May 2021*

- Relevant Coursework: Data Structures & C, Object Oriented Programming, Database Management System, Computer Networks, Information Theory, Mathematics, Wireless Communication and Radar Engineering
- Top Class Performer, Professor Recommendations, Final Year Project Recognition – *"Unmanned Sea Surveillance Boat with IoT"*

Projects

Automatic Skin Cancer Type Detection Using CNNs and Deep Learning | *Github*

- Developed a Sequential Convolution Neural Network (CNN) with the VGG16 model using TensorFlow and Keras to accurately classify skin cancer types from images, executed data pre-processing, pattern visualization, utilized data augmentation techniques to reduce overfitting by 15%, and verified evaluation metrics for consistent model performance.

WhatsApp Business Review Sentiment Analysis with Natural Language Processing | *Github*

- Integrated Recurrent Neural Networks (RNNs) with NLTK's VADER model and Part-of-Speech (POS) tagging to conduct sentiment analysis on customer reviews, providing in-depth insights into customer sentiments and supporting enhanced decision-making.

Statistical Analysis on Accidental Drug Deaths in Connecticut | *Website*

- Analyzed 51.3K entries in real-time using Python key-libraries to reveal trends in drug-related deaths in Connecticut between 2012 & 2022, employed advanced statistical techniques, EDA analysis, feature engineering and data visualizations.

Skills

Languages: Python, R, SAS, SQL, MATLAB, C++, Java, Bash, JavaScript, TypeScript, HTML & CSS
Libraries: Pandas, Numpy, MatplotLib, scikit-learn, SciPy, TensorFlow, Keras, OpenCV, Pytorch
Technologies: Flask, MongoDB, MS-Tools, Tableau, Power-BI, Snowflake, AWS, Azure, Grafana, Canva, Squarespace
Methodologies: ETL, Agile, Scrum, OOP, Functional Programming, DevOps, CI/CD