

Shreya Banik

☎ +1 (203) 676-7828 | ✉ shreyabanik2k@gmail.com | 🌐 [shreya-insights](https://shreya-insights.com) | 🐙 [shreya.github](https://github.com/shreya) | 📁 [shreya.portfolio](https://shreya.portfolio.com) | 📍 New Haven, CT

EXPERIENCE

University of New Haven

Connecticut, USA

Research Assistant

Jan 2025 – Present, Part-time

- Led a cross-functional team of 5 on a **deep learning-based 3D-Vision Welding Recognition system under SAIL-LAB**, focusing on **seam detection optimization** using **computer vision** ensuring timely project milestones and adherence to research goals.
- Evaluated the tensorflow-based **baseline model** and designed a custom large model **Attention U-Net** with **research-based architecture** and resolved cross-team blocker with the **GPU memory consumption** for large batch-size using CUDA Library.
- Conducted rigorous evaluation using metrics such as **validation loss**, **pixel-level Intersection over Union (IoU)**, **accuracy** and collected **good and hard** test samples; implemented a combined **BCE + Dice loss function** with coefficient tuning.
- Applied **deep learning optimization techniques** including 3 stages of **hyperparameter tuning**, **targeted fine-tuning**, and standardized **data augmentation & normalization strategies** to enhance model robustness on challenging welding data.
- Reported weekly deliverables through **interactive presentations using Canva**, and maintained comprehensive documentation of code, experiment results, and model evaluation reports on **GitHub** to ensure transparency and reproducibility.
- Collaborated with faculty and industry partners with to iteratively **test, refine, and validate the model** on diverse datasets, maintaining **data confidentiality** and ensuring **model generalization**.

WelSpot Inc.

New York, USA

Data Science Intern

Jan 2025 – Present, Intern

- Designed and deployed an **AI-powered business health monitoring and recommendation system** using **Vertex AI** and **Google Cloud Platform (GCP)**, enabling personalized, data-driven health-insights for strategic decision-making on WelSpot's platform.
- Facilitated **multi-stakeholder collaboration** across product, engineering, and business units to align on **system architecture**, document **machine learning requirements**, and deliver technical & strategic presentations comparing **generative AI Models**.
- Coordinated the end-to-end development of a **Retrieval-Augmented Generation (RAG) pipeline**, overseeing the integration of **GCP services**, **BigQuery**, and **Vertex AI** to deliver secure, scalable, and production-ready recommendation workflows.
- Engineered and optimized **large-scale data infrastructure** on **Google BigQuery** and **Looker Studio** by defining **real-time analytics pipelines** and optimizing schema design through advanced **SQL queries** for business intelligence reporting.
- Researched, implemented, and evaluated **machine learning models** for **recommendation engines**, aligning architecture selection with business goals around **performance**, **cost-efficiency**, **latency**, and **data privacy**.

Cognizant Technology Solution

Kolkata, India

Programmer Analyst

Oct 2021 – Jul 2023, Full-time

- Developed an **Automatic Reconciliation Tool** for an insurance client using **MySQL Workbench** and **MS Excel** to track and audit over **400,000 daily ServiceNow transactions**, resulting in a **78% reduction in manual effort and resource hours**.
- Identified and reduced root causes of recurring discrepancies with preventive measures, leading to an **8X reduction in defects**.
- Leveraged advanced **SQL queries** for data extraction and transformation, uncovering customer transaction anomalies (policy and claim related) and driving a **30% improvement in reporting accuracy**.
- Documented **policy and claims-related large datasets**, delivering **actionable insights** using **CANVA** presentation to technical and non-technical stakeholders.

CBNITS Private Limited

Kolkata, India

Data Analyst

Apr 2021 – Sep 2021, Intern

- Developed a secure and flexible mobile application for **cryptocurrency exchange and wallet storage**, enhancing functionality by **82%** through innovative features and collaborative problem-solving with cross-functional teams.
- Analyzed **large financial datasets** on **cryptocurrency exchange** to derive insights; designed a scalable **data pipeline** using **REST APIs**, **MySQL**, and **Microsoft OneDrive** for cloud-based storage.
- Collaborated with cross-departmental teams to solve application issues by maintaining **Excel reports** and implementing **innovative formulas**, boosting speed of operations 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, Leadership & Entrepreneurism
- Research Assistant, 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: Analog & Digital Electronics, Circuit Theory & Network LAB, Data Structures & C, Object Oriented Programming, 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"*

SKILLS

Programming Languages & Frameworks: Python, SQL, R, PySpark, Bash, Java, C++ , HTML/CSS, JavaScript, TypeScript, Flask

Machine Learning & AI: Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, LightGBM, Optuna, Ray Tune, OpenCV, Attention U-Net, Transfer Learning, CNNs, LSTMs, Transformers

Natural Language Processing (NLP): spaCy, NLTK, Hugging Face Transformers, Retrieval-Augmented Generation (RAG), SHAP, LIME, BERT, Text Classification, Entity Recognition, Sentiment Analysis

Computer Vision (CV): OpenCV, Image Augmentation, Semantic Segmentation, Object Detection, Attention Mechanisms, 3D Vision Image Processing

Data Engineering & Pipelines: BigQuery, Apache Spark, Microsoft Fabric, MySQL, MongoDB, SAP ECC Integration, SSIS, REST APIs, ETL Pipelines, Data Modeling, Schema Design, Data Quality Assurance

Cloud Platforms & MLOps: GCP (Vertex AI, BigQuery, Looker Studio), AWS, Azure, Docker, Render, GitHub Actions, CI/CD for ML Pipelines, Model Deployment & Monitoring

Data Analysis & Visualization: Power BI, Tableau, Looker Studio, Excel, Pandas, NumPy, Matplotlib, Seaborn, SciPy

Project Leadership & Collaboration: Agile Project Management, Stakeholder Communication, Technical Documentation, Cross-Functional Team Leadership, Problem Solving, Analytical Thinking, Adaptability, Continuous Learning

Domain Expertise: AI/ML in Healthcare, Financial Intelligence, Welding Defect Detection, Business Intelligence Systems, Research & Academic Collaboration

PROJECTS

Image Processing — 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.

Recurrent Neural Network — 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.

Time Series Forecasting & Anomaly Detection — Weather Data Prediction (TSA) | [Github](#)

- Developed a 5-Stage time-series forecasting framework for weather analysis, integrating trend analysis, seasonality detection, correlation exploration, anomaly identification, and ARIMA/SARIMA modeling, validated with stationarity tests and error diagnostics, delivered through a reproducible Python pipeline combining statistical and machine learning techniques.

Statistical Data Analysis — Connecticut Drug-related Accidental Death Records Analysis | [Website](#)

- Conducted an in-depth statistical analysis of 51.3K real-time entries on accidental drug deaths in Connecticut (2012–2022) using Python, leveraging EDA, feature engineering, and advanced statistical techniques to uncover critical trends and contributing factors, supporting public health insights.

IoT & Robotics — IoT-based Unmanned Sea Surveillance Vehicle (USV) | [GitHub](#)

- Developed an IoT-based unmanned sea surveillance vehicle (USV) for defense and rescue operations on sea on Arduino Circuit with DC-motor and Electronics component, integrating GPS for location tracking, a night-vision camera for object detection, and remote control via the Blynk app to enhance border security and disaster response.