

Shruti Houji

Data Scientist

sghouji@iu.edu • 812-778-4724
[GitHub](#) • [LinkedIn](#) • [Portfolio](#) Bloomington, IN 47408

Education

Master in data science

Indiana University, Bloomington, IN

Aug 2022 – May 2024

GPA - 3.762

Bachelor in computer engineering

Pune University, Pune, India

Jun 2016 – Apr 2020

GPA - 4

Technical Proficiencies

Programming Languages: C++, Python, R, Java, DAX

Data Analysis Tools: Microsoft Power BI, MS Excel VBA, Tableau, Matplotlib, Seaborn

Databases and Web Technologies: SQL, Microsoft SQL server, MongoDB, Neo4j, BigQuery, Snowflake, HTML5, JavaScript

Cloud Platforms: AWS (S3, EC2, Lambda, SageMaker), Google Cloud Platform

Other: Statistics, Data Structures, Tensorflow, Git, Flask, NLP, PySpark, Databricks, Pandas, NumPy, SciPy, Scikit-learn

Career Experience

Danfoss Power Solutions, Cleveland, OH

May 2023 - Aug 2023

Data Science Intern

- Interpreted complex datasets, performed quantitative analysis and exploratory data analysis to uncover trends, patterns, and, driving informed decision-making and strategic planning initiatives with a 5% increase in accuracy.
- Integrated diverse data sources into data warehouses, built data pipelines to extract, transform, and load (ETL) data and SQL databases enabling data modeling & processing to streamline workflow, reducing data processing time by 10%.
- Automated production planning reports using Python and Excel VBA macros, reducing creation time by 30 minutes.
- Crafted KPIs, 5 data visualizations, and productivity reports in Power BI, using Python scripts, and DAX language to deliver recurring Business Intelligence solutions, delivering data-driven insights and informing decisions.

Cognizant Technology Solutions, Pune, India

Dec 2020 – May 2022

Programmer Analyst - Developer

- Cleaned input files for data analytics, utilizing Python libraries and Microsoft Power BI to create graphical visualizations.
- Collaborated with a cross-functional team of 10 people on Jira tasks, developed code using .NET and Docker for text and barcode extraction to support analyzing imaging data and validation tests.
- Expedited input feature decision-making by 5%, streamlining process and enabling business requirements.
- Increased business process efficiency by 75% by implementation of OCR & software solutions to process financial data using Data cap application, .NET, Oracle, Python, agile methodologies with SDLC framework.

Key Projects

Skin Lesion Classification for melanoma detection

- Spearheaded an end-to-end skin cancer detection system leveraging machine learning and amazon services like S3, SageMaker, Lambda, IAM and API Gateway to streamline analysis achieving 10% faster processing.
- Trained a neural network model with MONAI Pytorch and AWS cloud infrastucture to reduce resource utilization by 20%.

Datawage Navigator

- Structured a robust MySQL database design allowing users to access job listings in specific locations within milliseconds.
- Integrated CRUD operations and dashboard page using Flask to get real time updates with 10% decrease in data latency.

Energy Analytics and Prediction in USA

- Employed time series models such as ARIMA, VAR, and LSTM for forecasting optimal energy resources for the upcoming decade by applying data driven approaches.
- Created a four-page dashboard using Flask, encompassing imports, exports, production, and consumption of the energy resource "Natural Gas" promoting sustainability.

Real-time Intrusion Detection Systems for IOT Networks using ML

- Built a dataset by self-generated IoT node using Tshark script to get instananeous network logs of around 1 GB data.
- Applied data mining techniques to develop an IDS using a Random Forest model with an accuracy of 99% with fronend notification to detect real-time attacks through websockets.

Home Credit Default Risk-Classification

- Engineered a neural network model in PyTorch and predictive models, addressing financial instability & enhancing loan lending processes, achieving a Kaggle score of 76%, and aiding in A/B Testing within diverse ecosystems.