

Shweta Mishra

Boston, MA • [Email](#) • [LinkedIn](#) • [GitHub](#) • [Portfolio](#)

Summary

Innovative Data Scientist with master's in data science, 4+ years of experience, promoted thrice for driving 80% reduction in returns credit loss & optimizing fleet size to minimize losses through predictive analytics and Generative AI. Actively seeking data science roles to leverage my advanced analytical capabilities.

Experience

- Squark, Remote, Machine Learning Engineer** 2024 – Present
- Developed **2+ Python** scripts using **H2O** for modeling 4 ML tasks (Classification, Regression, Univariate & Multivariate Time Series Forecasting), optimizing workflow for faster, scalable deployment on **AWS** with Docker & Kubernetes.
 - Created a **SHAP** report and leveraged LLMs for explainability of predictions, providing intuitive visualizations, enhancing **business understanding**, and reducing **query resolution time** by **5%**.
 - Evaluated and compared **GluonTS** library **forecasting models**, increasing forecasting performance by **9%** in terms of **RMSE** and **MAE**.
- Extern, Remote, Beats by Dre Consumer Insights Data Analyst** 2024 – Present
- Expedia Group Competitive Intelligence and Strategy Analyst**
- Executed **sentiment analysis** on **10,000+** consumer reviews using **Python**, **NLTK** & **SpaCy** libraries, integrated advanced natural language processing (NLP) with **Gemini AI** to derive actionable insights for product enhancements.
 - Analyzed **2** direct and indirect competitors in the **retail media network** sector, evaluated products, revenue models, and market share, presenting insights, guided **strategic planning** and enhanced competitive edge.
- Tenneco, Northville, MI, Global Supply Chain Data Scientist** 2023
- Leveraged SQL Server (**SSMS**) & **Python** to track **sales KPIs** and **order trends**, identified top **10** customers and well-performing sectors, increased deal closure rate by **6%**, and **reduced excess stock** by **20%**.
 - Designed **Power BI (DAX Queries)** dashboard, automated alerts, tracked vendors' annual returns, collaborated with **cross-functional** teams, monthly data refresh via **Power Automate**, **boosted returns** management by **40%**.
- Auxi, Remote, Machine Learning Engineer** 2023
- Automated content creation for PowerPoint decks by integrating **OpenAI's LLMs** for Gen AI content development in **Flask App**, **increased product satisfaction** score by **10%** in a month.
 - Performed exploratory data analysis (EDA) using **Pandas**, **NumPy**, and **Matplotlib**, identified **5** key trends and patterns in **large datasets**, and presented findings with **Tableau** to inform strategic brand-building initiatives.
- Accenture, Bengaluru, IND, Sr Software Engineer – Finance & Supply Chain Team** 2018 – 2021
- Crafted a dynamic reconciliation report in **Power BI**, sourced using **advanced SQL**, aligned Project Costing with the General Ledger, ensured precise tallies before **monthly ledger closing**, eliminated **30%** of discrepancies.
 - Architected a robust predictive model using Random Forest and **XGBoost** algorithms, leveraged **scikit-learn** & **TensorFlow** framework to forecast **optimal fleet size** requirements, **improved fleet utilization** by **34%**.

Projects

- PDF QA using RAG • Langchain, Hugging Face, LLMs, Streamlit** 2024
- Engineered document-based interactive query tool for PDFs, utilized RAG pipeline, harnessed **Langchain** framework, **Hugging Face** embeddings, **FAISS** vector-store, & LLMs (**OpenAI**), **92%** reduction in processing time.
- Multimodal Forecasting of Solar Irradiance • ARIMA, Prophet, LSTM, Neural Nets, Transfer Learning** 2023
- Forecasted time-series data of solar irradiance, used **Inception V4** for sky images & **Prophet** for weather data, used Hugging Face model & **PyTorch** framework, achieved **MAE** of **74**.

Education

- Northeastern University, Boston, MA, MS in Data Science.** 2023
- KIIT University, Bhubaneswar, IND, B. Tech in Electronics & Telecommunication.** 2018

Skills & Interests

Technical: Data Wrangling, Statistical (**statsmodels**)/Predictive Analysis, Time-series Forecast, Plotly Dash, Dashboarding.
Language & Tools: SQL, Python, R, Excel, dbt, **PySpark**, GCS, BigQuery, Looker, Databricks, Git, Jira, Jenkins, Confluence.
Certification: [Microsoft Azure Certified Associate Data Scientist \(Synapse Analytics, Azure ML Studio\)](#)
Fun Fact: Hackathons Participant ([Hack Harvard 2023](#), [HACKATRON 2022](#)), Hiking, Odissi Dancer, Yoga Advocate.