

DISHA LAMBA

+1 (929) 444-1550 | jb.dishalamba@gmail.com | dldisha.github.io | [GitHub](#) | [LinkedIn](#)

PROFESSIONAL SUMMARY

Results-driven Data Scientist with 2 years of experience designing and deploying end-to-end machine learning pipelines and conducting large-scale data analysis. Proven expertise in EdTech, Healthcare, and consulting, with a track record of delivering insights that drive business growth. Skilled at aligning technical solutions with stakeholder needs through clear communication.

WORK EXPERIENCE

CVS Health

New York, NY

Data Scientist

Nov 2024 – Present

- Designed an LLM-driven note summarization system to extract key insights from unstructured discharge notes, retaining accurate critical information, improving care managers' efficiency and patient outreach across a 1.2M-member population.
- Optimized prompt engineering workflows, testing LLM architectures and refining prompts to enhance summary coherence.
- Designed KPIs and built Tableau dashboards to track usage, providing visibility that informed future feature development
- Collaborated with engineers, clinicians, & product teams to integrate the model into workflow and streamline operations.

Curate Partners (CVS Health)

New York, NY

Data Scientist

Aug 2024 – Nov 2024

- Led end-to-end data preparation of a 27M-record healthcare dataset to create a comprehensive patient encounter discharge notes data repository for downstream ML applications, ensuring HIPAA and FHIR compliance.
- Engineered complex BigQuery SQL pipelines to join and transform clinical data across multiple source systems.
- Conducted EDA using Python to identify discharge pattern anomalies, patient flow bottlenecks, and key clinical insights.
- Developed documentation for data schemas, processing workflows, and key insights to facilitate cross-team collaboration.

Mindgrasp

Columbia, MD

Full-Stack Machine Learning Engineer

Jan 2024 – April 2024.

- Designed and deployed 'Math QnA' — a serverless, LLM-powered feature that enabled users to solve complex math queries instantly, resulting in a 12% increase in user growth through improved engagement and functionality.
- Architected the solution using AWS Lambda for backend computation, AWS API Gateway for secure, low-latency communication, and AWS S3 for structured result caching. Integrated the system seamlessly into a ReactJS frontend, achieving sub-second response times.
- Conducted user testing and used Mixpanel to analyze user metrics, delivering insights that guided product development.

SKILLS

Languages & Packages: Python, SQL, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Keras, TensorFlow, HTML/CSS, ReactJS

Machine Learning: NLP, Regression, Classification, Clustering, Statistical Analysis, Time-Series Analysis, Neural Networks

Database & Management: Data Warehousing, Data Modelling, MySQL, PostgreSQL, Redshift, BigQuery, Spark, Airflow

Cloud & Other Tools: Amazon Web Services, Google Cloud Services, Git, Tableau, Mixpanel, Jira, Confluence, MS Office Suite

EDUCATION

New York University

New York, NY

Master of Science in Computer Engineering

May 2023

Coursework: Machine Learning, Deep Learning, Statistics, Big Data, Decision Optimization Models, Data Analytics

Guru Gobind Singh Indraprastha University

Delhi, India

Bachelor of Technology in Information Technology

September 2020

Coursework: Data Structures, Computer Architecture, Cloud and Mobile Computing, Software Engineering, DBMS

CERTIFICATIONS & ACHIEVEMENTS

- [IBM Data Science Professional Certificate](#), Coursera – June'24
- Best Medical Hack** at [MHacks 13 hackathon](#); developed ReactJS app to match COVID-19 patients with plasma donors.