DISHA LAMBA

+1 (929) 444-1550 | jb.dishalamba@gmail.com | dldisha.github.io | linkedin.com/in/dldisha/

PROFESSIONAL SUMMARY

Experienced Data Scientist with 3 years of experience building scalable ML systems and LLM applications that drive measurable business impact. Proven track record architecting end-to-end AI solutions from data infrastructure to production deployment. Strong foundation in Python, SQL, Tableau, statistical analysis, experimentation, and translating technical insights for C-level stakeholders.

WORK EXPERIENCE

CVS Health   New York, NY
Data Scientist   Aug 2024 – Present
- Achieved **80% reduction in clinicians' patient-note review time** by architecting an **end-to-end LLM system** that extracts key insights from unstructured discharge notes for **1.2M-member population**, implementing **schema-guided prompting with Pydantic** and **ClinicalBERT-based hallucination detection**, with continuous evaluation metrics to ensure clinical accuracy.
- Built **LLM-based risk detection model** on longitudinal claims data, enabling early chronic-disease interventions with **~$100M projected savings**. Developed **Streamlit tool that automated prompt generation** through templating, **reducing manual effort by ~60%** and accelerating post-model-design experimentation. Designed evaluation metrics and ran validation experiments with clinicians to iteratively refine detection criteria and ensure clinical validity before production deployment.
- Led **CVS Specialty Pharmacy** project benchmarking 36 pharmacies' data-sharing practices to guide strategic decisions.
- **Engineered 20M dataset** using SQL by unifying multiple data sources, **reducing access time by 70% for downstream ML work**.
- Designed and built **20+ SQL-based behavioral and engagement metrics** tracking user patterns and efficiency trends, integrating them into **Tableau dashboards** to give stakeholders real-time visibility and support data-informed decisions.
- **Mentored 3 junior associates** on **ML best practices**, fostering technical growth and strengthening team capabilities.
- **Presented insights** to **C-level** and cross-functional stakeholders, enabling adoption of **data-driven decisions across workflows**.

Mindgrasp   Columbia, MD
Machine Learning Engineer   Jan 2024 – April 2024
- Designed a **math solver using Wolfram Alpha API**, enabling users to solve math problems, **increasing user retention by 18%**.
- Architected scalable serverless solution with **AWS Lambda, API Gateway, and S3 caching**, achieving <500ms latency and 99.9% uptime while handling 10K+ daily requests through ReactJS frontend integration, with **CloudWatch monitoring**.
- Ran **A/B experiments** with **Mixpanel** to **analyze user behavior** and feature adoption, driving data-informed iterations.

SKILLS

Languages & Packages: Python, SQL, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Keras, TensorFlow, HTML/CSS, ReactJS
Machine Learning: LLMs, RAG, Prompt Engineering, NLP, Regression, Classification, Clustering, A/B Testing, Statistical Analysis, Time-Series Analysis, Recommendation Systems
Database & Management: Data Warehousing, Data Modelling, MySQL, PostgreSQL, Redshift, BigQuery, Spark, Airflow
Cloud & Other Tools: Amazon Web Services, Google Cloud Services, Docker, Tableau, Mixpanel, Jira, Confluence, MS Office Suite

PROJECTS

Recommendation System using MovieLens Dataset
- Built a scalable movie recommendation system on **MovieLens 32M dataset using PySpark's ALS algorithm** for latent factor modeling, benchmarking against user-based and popularity-based approaches to maximize accuracy.

Adaptive Email Agent System using DSPy
- Developed AI agent using DSPy to generate personalized email responses by analyzing context and self-optimizing prompts.

EDUCATION

Master of Science in Computer Engineering   New York, USA
New York University   May 2023
Coursework: Machine Learning, Deep Learning, Probability & Statistics, Big Data, Decision Optimization Models, Data Analytics

Bachelor of Technology in Information Technology   Delhi, India
Guru Gobind Singh Indraprastha University   September 2020

ACHIEVEMENTS

- **Speaker**, AICamp NYC 2024, Data Summit Boston 2025
- **Best Medical Hack** at MHacks 13 hackathon; developed ReactJS app to match COVID-19 patients with plasma donors.