# DISHA LAMBA

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

## PROFESSIONAL SUMMARY

Results-driven Data Scientist with 3 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**
Clinical Data Scientist                                                                                      August 2024 – Present
- Developed a note summarization system leveraging LLMs to condense lengthy discharge notes into 1-page JSON summaries, assisting care managers in reducing hospital readmissions and enhancing patient outreach.
- Optimized data pipelines by designing SQL queries in BigQuery to process 27M dataset, ensuring HIPAA compliance.
- Collaborated with engineers and nurses to refine the model, integrating insights for enhanced workflow efficiency.

**Mindgrasp**                                                                                                         **Columbia, MD**
Data Scientist                                                                                                  Jan 2024 – April 2024
- Developed Math QnA, a feature utilizing Gen AI and AWS Lambda to process advanced math queries, managing data with AWS S3. This solution served over 1M users and contributed to a 12% increase in user growth.
- Led a team of 5 developers in improving ReactJS app performance through A/B testing, enhancing user experience.
- Utilized Mixpanel to analyze user metrics, driving data-informed product development, & delivered key insights to C-suite

**Subconscious AI**                                                                                                   **New York, NY**
Machine Learning Engineer                                                                              July 2023 – Jan 2024
- Optimized the research tool by preprocessing data and applying decision trees with Python, SQL achieving 83% accuracy.
- Integrated Hate-Speech detection API using TensorFlow, improving moderation accuracy by 90% and saving 200+ hours.
- Streamlined project timelines & resource utilization in Agile, Git-based environment, enhancing development lifecycle.

**Data Glacier**                                                                                                      **New York, NY**
Data Science Intern                                                                                       June 2022 – Aug 2022
- Engineered a Python-based BERT model to develop a semantic text-ranking search system that achieved 92% accuracy.
- Leveraged data analytics and ETL pipelines with SQL and Airflow to extract insights from 1M documents for data usability.
- Optimized data storage, processing using AWS Lambda, S3 achieving 23% improvement in data access and cost reduction.
- Collaborated with product team and presented insights to stakeholders via Tableau, PowerPoint, driving business growth.

## 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.