

# DISHA LAMBA

+1 (929) 444-1550 | [dishalamba01@gmail.com](mailto:dishalamba01@gmail.com) | [sudodisha.github.io](https://sudodisha.github.io) | [GitHub](#) | [LinkedIn](#)

## PROFESSIONAL SUMMARY

Accomplished Data Scientist with **3+** years of experience demonstrating deep understanding of end-to-end machine learning pipelines, predictive modeling, large dataset analysis, and stakeholder communication. Proven track record of successfully optimizing business problems, answering critical business questions, and contributing to mission-critical decision-making.

## WORK EXPERIENCE

- UnitedHealth Group**  
Data Scientist

**New York, NY**  
June 2023 – Present

  - Led cross-functional projects improving medical practices, reducing readmissions and opioid use, delivering \$5M savings.
  - Employed advanced SQL queries for data cleaning, using subqueries, window functions, to extract trends and anomalies.
  - Developed a predictive logistic regression model, leveraging Pandas for exploratory analysis and data structuring, and Scikit-learn and Matplotlib for model building, to forecast patient hospitalization using medical history data.
  - Conducted A/B testing for user behavior analysis and segmentation, presenting insights to stakeholders using Tableau reports and dashboards to inform business strategies and enable informed decision-making.
  - Implemented CI/CD practices and used Git for version control, ensuring seamless collaboration and efficient deployment.
  - Worked on Cloud Services - AWS S3, Redshift to assist with big data tools, solve storage issues, and improve scalability.
- Data Glacier**  
Data Science Intern

**New York, NY**  
June 2022 – Aug 2022

  - Developed semantic text-ranking search system using Hugging Face Transformers to enhance document retrieval accuracy, achieving 42% reduction in query processing time and enhancing search result accuracy.
  - Engineered Python-based BERT model, achieving 92% accuracy in semantic understanding and search result relevance.
  - Built scalable ETL pipelines using Airflow to process 1M documents, improving data usability and supporting data analytics.
  - Optimized data storage, processing using AWS Lambda, S3 achieving 23% improvement in data access and cost reduction.
  - Applied data visualization techniques to present findings, resulting in actionable insights and strategic decisions.
- Accenture**  
Data Scientist

**Haryana, India**  
Jan 2020 – June 2021

  - Leveraged advanced data analytical techniques and ETL pipelines to extract actionable insights from complex datasets, contributing to improved operational efficiency and strategic decision-making.
  - Developed XGBoost classification model on AWS SageMaker, leveraging AWS S3 for scalable financial data storage, utilizing Python, NumPy, and hyperparameter tuning to predict credit payment delays, increasing approval rates by 11.5%.
  - Enhanced model accuracy and reliability through EDA using Pandas, Matplotlib, followed by feature selection techniques, reducing financial risks and improving lending decisions, directly saving ~\$3M in potential losses.
  - Contributed to product roadmap development and assisted in backlog prioritization, enhancing team efficiency via Jira.
  - Applied statistical and data visualization techniques with Tableau for actionable insights, and data-driven decision-making.

## SKILLS

**Languages & Packages:** Python, R, SQL, Pandas, NumPy, Matplotlib, Plotly, 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, Spark, Kafka, Airflow  
**Cloud & Other Tools:** Amazon Web Services, Terraform, Git, Linux, MS Excel, Tableau, Power BI, Mixpanel

## EDUCATION

- New York University**  
Master of Science in Computer Engineering  
Coursework: Machine Learning, Deep Learning, Statistics, Big Data, Decision Optimization Models, Data Analytics

**New York, NY**  
May 2023
- Guru Gobind Singh Indraprastha University**  
Bachelor of Technology in Information Technology  
Coursework: Data Structures, Computer Architecture, Cloud and Mobile Computing, Software Engineering, DBMS

**Delhi, India**  
September 2020

## 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.