

DISHA LAMBA

+1 (929) 444-1550 | dl4747@nyu.edu | sudodisha.github.io | [GitHub](#) | [LinkedIn](#)

PROFESSIONAL SUMMARY

Experienced Data Scientist with 3 years of experience leveraging statistical methods and machine learning to build predictive models informing data-driven strategies. Proficient in Python, SQL, and AWS, with strong ability to extract insights from complex datasets and deliver clear narratives. Passionate about new technology and actively involved in developer communities.

WORK EXPERIENCE

- | | |
|--|---|
| UnitedHealth Group
Data Scientist | New York, NY
June 2023 – Present |
| <ul style="list-style-type: none">Led 2 collaborative projects with marketing, legal, medical, and engineering teams, driving medical behavior changes that decreased hospital readmissions and opioid use, saving ~\$5M annually.Developed logistic regression model in Python, utilizing Pandas, SQL for data manipulation and Scikit-learn for model building, to forecast patient hospitalization and opioid use disorder risks, using health and medical history data.Conducted A/B tests to assess communication strategy effectiveness, monitoring KPIs like behavior change, cost savings.Presented insights to stakeholders using Tableau dashboards, enabling informed decision-making and strategic planning. | |
| Data Glacier
Data Science Intern | New York, NY
June 2022 – Aug 2022 |
| <ul style="list-style-type: none">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.Orchestrated ETL pipelines to process 1M unstructured documents into structured data with Airflow, enhancing usability.Optimized data storage, processing using AWS Lambda, S3 achieving 23% improvement in data access and cost reduction. | |
| Accenture
Data Scientist | Haryana, India
Jan 2020 – June 2021 |
| <ul style="list-style-type: none">Developed a classification XGBoost model on AWS SageMaker, leveraging AWS S3 bucket for secure, scalable storage of financial data 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.Presented strategic model insights to senior leadership, led multiple data analysis projects, and managed cross-functional teams using Agile methodologies, optimizing workflow efficiencies through Jira.Spearheaded development of scalable data infrastructure that improved code collaboration with Git version control.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, MySQL, PostgreSQL, AWS S3, AWS Redshift, Spark, Kafka, Airflow, dbt
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
Course Assistant: Deep Learning, Computer Networking | 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
- AWS Certified Cloud Practitioner** – June'24
- Conference Grant recipient** for AWS re:Invent'23 and Strange Loop'22.
- Best Medical Hack** at [MHacks 13 hackathon](#); developed ReactJS app to match COVID-19 patients with plasma donors.