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 data analysis and machine learning to inform product strategies. Proficient in Python, SQL, and AWS, extracting insights from complex datasets and delivering clear narratives. Skilled in cross-functional teamwork to drive product development. Passionate about technology and active in developer communities.

WORK EXPERIENCE

UnitedHealth Group   **New York, NY**
Data Scientist   June 2023 – Present
- Led cross-collaborative, data-driven projects that reduced hospital readmissions and opioid use, saving ~$5M.
- Employed advanced SQL queries for data cleaning, using subqueries, window functions, to extract trends and anomalies.
- Developed predictive logistic regression model, leveraging Pandas for data structuring, EDA and Scikit-learn, Matplotlib for model building, to forecast patient hospitalization and opioid use disorder risks, using 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   **New York, NY**
Data Science Intern   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.
- 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   **Haryana, India**
Data Scientist   Jan 2020 – June 2021
- 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.
- Spearheaded development of scalable data infrastructure that improved code collaboration with Git version control.
- Contributed to product roadmap development, participated in sprint planning sessions, and assisted in backlog prioritization using data-driven insights, enhancing team efficiency through 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, AWS S3, AWS Redshift, Spark, Airflow, dbt
Cloud & Other Tools: Amazon Web Services, Git, Linux, Tableau, Power BI, Mixpanel, Jira, Asana, Microsoft 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
Course Assistant: Deep Learning, Computer Networking

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
- **AWS Certified Cloud Practitioner** – June'24
- **Former Vice President** at NYU Machine Learning Club'23
- **Best Medical Hack** at MHacks 13 hackathon; developed ReactJS app to match COVID-19 patients with plasma donors.