

DISHA LAMBA

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

EDUCATION

New York University

Master of Science in Computer Engineering

Coursework: Machine Learning, Deep Learning, Statistics, Big Data, Decision Optimization Models and Data Analytics

Course Assistant: Deep Learning, Computer Networking

NY, USA

May 2023

Guru Gobind Singh Indraprastha University

Bachelor of Technology in Information Technology

Coursework: Data Structures and Algorithms, Cloud Computing, Software Engineering, Database Management Systems

Delhi, India

September 2020

SKILLS

Technical: Python, R, SQL, SAS, Postgres, Hadoop, Spark, Kafka, PowerBI, Tableau, MS Excel, Monte Carlo Simulation, TensorFlow, SciPy, NumPy, Pandas, Keras, Scikit-learn, Matplotlib, ggplot, HTML, CSS, ReactJS, Flask

Cloud Services and Tools: Amazon Web Services (AWS), Terraform, Git, Linux, Docker

Others: Data Mining, Regression, Classification, Clustering, Statistical Analysis, Time-Series Analysis, A/B Testing

WORK EXPERIENCE

Mindgrasp

Data Engineer Intern

- Enhanced user interaction analysis with PostHog, Mixpanel, and A/B testing, for efficient tracking of user journeys and conversion rates.

Columbia, MD

Jan 2024 – Present

Subconscious AI

Machine Learning Developer Intern

- Optimized Market Research tool using Python and SQL, through data pre-processing, and feature engineering; applied decision trees, achieving 30% improvement in data quality and 83% model accuracy, enhancing predictive insights.
- Implemented content filtering using Hate-Speech API, achieving 90% moderation improvement, saving 200+ hours.
- Streamlined project timelines and resource utilization in a Git and Agile environment, enhancing the development lifecycle.

New York City, NY

July 2023 – Jan 2024

Data Glacier

Data Science Intern

- Developed a semantic text-ranking search engine, streamlining search processes and achieving 45%-time efficiency gain.
- Implemented ETL and NLP methods to convert unstructured text data into structured formats for efficient data usability.
- Attained 92% accuracy using Python-based BERT and TF-IDF ensemble model, enhancing search capabilities.
- Optimized data storage & processing with AWS S3, Lambda, achieving 25% data access improvement, and cost reduction.

New York City, NY

June 2022 – Aug 2022

Sapio Analytics

Machine Learning Developer

- Spearheaded development of job-search portal for the Government of India, generating 1.2M job opportunities nationwide.
- Achieved 85% accuracy in user-job matching using sklearn K-nearest neighbor algo, tailored for user location, skillset.
- Collaborated with software developers on React D3.js for data visualization, resulting in 25% increase in user engagement.
- Devised A/B experiments in collaboration with product and business team for data-driven decision-making and testing.
- Conceptualized and presented key results to senior leadership using Tableau, facilitating strategy development.

Mumbai, India

July 2020 – June 2021

PROJECTS

Data Analytics - Portfolio Optimization (MS Excel, Tableau)

- Revamped deterministic model with Excel Solver for portfolio optimization, resulting in 18% decrease in projected costs.
- Developed Tableau dashboards for portfolio performance, risk metrics, KPIs for informed investment decision-making.

Big Data - Movie Recommendation System (Python, PySpark, SQL) [[GitHub](#)]

- Extracted valuable insights from 27M MovieLens dataset using SQL, Pandas, Matplotlib guiding feature development.
- Engineered collaborative filtering ALS algorithm in PySpark, benchmarking against Python LightFM and Annoy models, achieving 95% accuracy with Annoy for top 3 movie recommendations, continuously adapting to user preferences.

ACHIEVEMENTS AND LEADERSHIP

- Best Medical Hack** at [MHacks 13 hackathon](#); developed ReactJS app to match COVID-19 patients with plasma donors.
- Conference attendee** at AWS re:Invent'23 (Scholar), Grace Hopper Celebration'23, Strange Loop'22 (Scholar).
- Former **Vice President**, NYU Machine Learning Club: Ignited machine learning interest in 30+ students.