

DISHA LAMBA

(929) 444-1550 | dl4747@nyu.edu | <https://dldisha.github.io> | [GitHub](#) | [LinkedIn](#)

EDUCATION

New York University, New York, NY

Sept 2021 - May 2023

MSc Computer Engineering

GPA: 3.5

Coursework: Machine Learning, Deep Learning, Big Data, Decision Models and Analytics, Network Security

Bharati Vidyapeeth's College of Engineering, New Delhi, IN

Aug 2016 - Sept 2020

Bachelor's of Technology Information Technology

GPA: 3.4

Coursework: Data Structures and Algorithm, Operating System, Compiler Design, Software Engineering, Web Development

TECHNICAL SKILLS

Languages

Python, R, C/C++, JavaScript, Git, Linux, HTML/CSS

Packages and Libraries

NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, PyTorch, TensorFlow, Keras, dplyr, NLTK

Tools & Frameworks

Flask, React.js, Ruby on Rails, Tableau, MS Excel, AWS, GCP, Kanban

Databases

SQL, PostgreSQL, Hadoop, PySpark, MongoDB

PROFESSIONAL EXPERIENCE

New York University

New York, NY

Course Assistant

Sept 2022 - May 2023

- Designated as Graduate Teaching Assistant for the Deep Learning Fall'22 and Computer Networking Spring'23 courses.
- Assisted more than 100 students with lectures, assignment queries via weekly office hours, in addition, to grading assignments.

Data Glacier

New York, NY (Remote)

Data Science Intern

June 2022 - Aug 2022

- Developed a search engine that ranks unstructured text documents based on semantic & contextual relationships to user prompts.
- Mined unstructured text data of **1M incidents in SQL** server to generate Document-Term-Matrix with TF-IDF weights.
- Proposed and built an ensemble of BERT and TF-IDF, capturing **80% of escalations** and achieving a **MAP score of 80.3%**.

Sapio Analytics

Maharashtra, IN (Remote)

Machine Learning Engineer

July 2020 - June 2021

- Spearheaded development of a job-search web app for the **Govt. of India**, resulting in **1.2M blue-collar jobs** to Indian laborers.
- Analyzed 27M** of supply chain data, generated DQIs using Informatica, and built changes to enhance the data quality by 17%.
- Designed & Developed a job-search platform **using TensorFlow** that utilized KNN algorithm to match laborers' skill sets with industry requirements; **generated a 50%+ increase in job placements** for unemployed blue-collar workers during Covid.
- Incorporated the KNN model into Elasticsearch to enhance job matching accuracy, resulting in an **85% accuracy rate**.
- Collaborated with the SDE team to seamlessly integrate ML model with frontend using React to display job matches effectively.

PROJECTS

Finance Portfolio Management System (Flask, SQL)

Sept 2022 - Dec 2022

- Developed a comprehensive financial portfolio management system using Flask, allowing users to track and analyze their financial activities in a secure and user-friendly manner including features such as expense tracking, budgeting, and goal setting.
- Integrated third-party APIs to fetch real-time financial data, enabling users to view account balances, transaction histories, etc.

Movie Recommendation System (Python, PySpark)

Jan 2022 - May 2022

- Analyzed & processed 27M records** of movie-user information in PySpark, and leveraged Python-Seaborn to plot graphs.
- Implemented collaborative filtering movie recommendation model using Spark's ALS algorithm to suggest top 10 movies based on genres. Conducted a comparative analysis of the baseline model's performance with Python libraries LightFM and Annoy.
- Achieved superior results using LightFM over ALS method, with an **MSE score of 0.77** and computational **time of 95.12 sec**.

Dialogue-based Interactive Facial Editing via Reinforcement Learning (PyTorch, NLP)

Jan 2022 - May 2022

- Implemented a dialogue-based facial editing system using reinforcement learning that allows users to specify the desired facial editing effect in a natural language and interactively accomplish the goal via multimodal dialogue with system feedback.
- Devised quantitative & qualitative experiments to validate system performance, with 80% of participants favoring model results.

PUBLICATION

Paper titled '**An Integrated System for Occupational Category Classification based on Resume and Job Matching**' published in *International Conference of Data Analytics & Management*, Springer, June 2020 | [Paper link](#)

LEADERSHIP

- Vice President at NYU Machine Learning Club

Jan 2023