

DISHA LAMBA

New York, USA

dl4747@nyu.edu | <https://dldisha.github.io> | github.com/dldisha | [linkedin.com/in/dldisha/](https://www.linkedin.com/in/dldisha/)

EDUCATION

New York University | New York, USA

Sept 2021 - May 2023

MSc Computer Engineering

GPA: 3.5

Relevant courses: Machine learning, Deep learning, Big data

Bharati Vidyapeeth's College of Engineering | Delhi, India

Aug 2016 - Sept 2020

Btech Information Technology

GPA: 3.3

TECHNICAL SKILLS

Languages: Python, R, JavaScript, Git, Linux, HTML/CSS

Tools and Frameworks: NumPy, Pandas, Scikit-learn, Seaborn, Matplotlib, Tableau, PyTorch, TensorFlow, React.js, Ruby on Rails

Databases: MySQL, PostgreSQL, DynamoDB, MapReduce, PySpark

PROFESSIONAL EXPERIENCE

Data Glacier | Data Science Intern | New York (Remote)

July 2022 - Present

- Developed a search engine that ranks unstructured text documents based on semantic & contextual relationships to user prompts.
- Proposed and built an ensemble of BERT & TF-IDF on MS Macro dataset, achieving MAP of 82.3%.

Sapio Analytics | Machine Learning Intern | Maharashtra, India (Remote)

July 2020 - Jan 2021

- Designed a deep learning model using TensorFlow that maps the skill set of laborers with those of industry requirements across Pan-India to help them with blue-collar jobs during covid. Deep-KNN was used as a matching algorithm.
- Extracted and analyzed specific data for both industry-labor demand-supply using Python, SQL to understand industry demand.
- Developed and analyzed Tableau reports and dashboards of over 20M records using PySpark ETL scripts and hive tables.
- Integrated TensorFlow model with frontend using ReactJs to display state-wise job requirements, industry analysis, and matches.
- Initiated web application for the Govt. of India to provide 1M blue-collar jobs to Indian migrant laborers.

inDDev | Software Developer Intern | Haryana, India

June 2019 - July 2019

- Designed and implemented the frontend and backend architecture of a Content Management System(CMS) using Ruby on Rails.
- Implemented version-controlled environment in CMS; allows the admin to safeguard any changes and roll back when necessary.
- Extended the web application with built-in ruby gems making it SEO-friendly with ancestry, and drag and drop functionality.

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**)

PROJECTS

Movie Recommendation System using PySpark

- Studied over 27 million records of movie-user information, and leveraged Tableau to plot graphs, and clean data in PySpark.
- Implemented a 'collaborative filter-based' movie recommendation system using Spark's ALS method, comparing it with that of a baseline model as well as with single-machine implementation (LightFM) and Annoy(Approximate Nearest Neighbors).
- Tested each model for overall performance using accuracy metrics like precision at k, NDCG, MAP, and RMSE.
- Single-machine implementation performed better than the ALS method with a test MAP of 0.057 and 145.12sec to fit the model.

Dialogue-based Interactive Facial Editing via Reinforcement Learning

- Developed 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.
- To track user intention and maximize user's long-term rewards; employed policy gradient-based algorithm REINFORCE.
- Proposed a new user simulator with human evaluation to verify if the simulator can successfully simulate real user behaviors.

Plasma Desk web app - ExpressJs

- Developed a web platform, driven by doctors, that connects an eligible plasma donor with a Covid patient, based on compatibility.
- Web portal was developed with Firebase as its backend and served using ExpressJs and Javascript.
- The project won the **Best Medical Hack** and **Wolfram Award for Top 30 hacks** in MHacks 13 Beta Hackathon 2020.

ACHIEVEMENTS

- 3rd position in CSAW'21** Cyber Security Games & Conference Hack3d competition. Nov 2021
- Top 100/2000 Hacks in Facebook Messaging Hackathon 2020.** Aug 2020