# DISHA LAMBA

New York, NY

(929) 444-1550 | dl4747@nyu.edu | https://dldisha.github.io | github.com/dldisha | linkedin.com/in/dldisha/

## 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

## PROFESSIONAL EXPERIENCE

**New York University**                                                              New York, NY
**Course Assistant**                                                            Sept 2022 - Present
- Designated as Graduate Teaching Assistant for the Deep Learning Fall'22 and Computer Networking Spring'23 courses.
- Assisting 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 text documents in SQL server to generate Document-Term-Matrix with TF-IDF weights.
- Employed an ensemble of BERT and TF-IDF, capturing ~80% of escalations achieving a MAP of 80.3%.

**Sapio Analytics**                                                        Maharashtra, IN (Remote)
**Machine Learning Engineer**                                                   July 2020 - June 2021
- Initiated and launched the web application for the Govt. of India to provide ~1.2M blue-collar jobs to Indian migrant laborers.
- Designed and built a Job-Search Portal 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. KNN was used as a matching algorithm.
- Analyzed Tableau reports and dashboards of over ~20M records using PySpark ETL scripts to understand the supply chain.
- Integrated the KNN model with ElasticSearch to match the job and user requirement; achieved an accuracy of 85%.
- Collaborated with the software team to integrate TensorFlow model with frontend using React to display analysis and matches.

## 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** | React.js, Ruby on Rails, Tableau, MS Excel, Amazon Web Services(AWS) |
| **Databases** | SQL, PostgreSQL, Hadoop, PySpark, DynamoDB |

## PROJECTS

**Movie Recommendation System (Python, PySpark, SQL)**                      Jan 2022 - May 2022
- Studied over 27M records of movie-user information, clean data in PySpark, and leveraged Python-Seaborn to plot graphs.
- Implemented a baseline collaborative filtering movie recommendation model using Spark's ALS method to recommend the top 10 movies based on movie genres, implemented and compared baseline model results with Python libraries LightFM and Annoy.
- LightFM performed better than the ALS method with test scores - MSE of 0.77, AUC of 0.88, and 95.12sec to fit the model.

**Dialogue-based Interactive Facial Editing via Reinforcement Learning (Python, PyTorch, NLP)**   Jan 2022 - May 2022
- 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.

**Plasma Desk web app (ExpressJs, HTML/CSS, Firebase, AWS - ElasticSearch, Heroku)**             Aug 2020
- Developed JS web app, driven by doctors, that connects an eligible plasma donor with a Covid patient, based on compatibility.
- The project won the **Best Medical Hack** and **Wolfram Award for Top 30 hacks** in MHacks 13 Beta Hackathon 2020.

## 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)**

## ACHIEVEMENTS

- **Vice President** at **NYU Machine Learning Club**                                     Jan 2023
- **100% Grant Recipient** for attending **Strange Loop Conference 2022** in St. Louis, Missouri    Sept 2022
- **3rd position in CSAW 2021** Cyber Security Games & Conference Hack3D competition.         Nov 2021