# DISHA LAMBA

New York, USA

dl4747@nyu.edu | https://dldisha.github.io | github.com/dldisha | linkedin.com/in/dldisha/

## TECHNICAL SKILLS

**Languages:** Git, Linux, JavaScript, R, Python, C/C++, HTML/CSS
**Tools and Frameworks:** Ruby on Rails, React.js, PyTorch, TensorFlow, RStudio, Microsoft Excel, Tableau
**Machine Learning:** Regression, Feature Selection, Clustering, Decision Trees
**Databases:** MySQL, PostgreSQL, DynamoDb, MapReduce, PySpark

## PROFESSIONAL EXPERIENCE

**Sapio Analytics | Machine Learning Intern** | *Maharashtra, India (Remote)* July 2020 - Jan 2021

- Designed, trained and deployed a **predictive machine learning model** using TensorFlow that maps the skill set of users with those of industry requirements to help migrant labores across Pan-India with blue-collar jobs during the time of covid.
- 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 over large datasets using **PySpark** ETL scripts and hive tables.
- Integrated machine learning model with frontend using **ReactJs** to display state-wise job requirements and each industry analysis.
- The web application is being used by **Govt. of India** and is able to provide **1 million blue-collar jobs** to Indian laborers.

**inDDev | Software Developer Intern** | *Haryana, India* June 2019 - July 2019

- Designed and implemented the frontend and backend architecture of a **Content Management System** using **Ruby on Rails**.
- Implemented a **version-controlled environment** that allows the admin to safeguard any changes and roll back when necessary.
- Extended the 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, 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)**.
- Conducted **A/B testing** experiments, performed statistical analysis for sample size requirements and experiment duration.
- Tested each model for accuracy metrics like precision at k, NDCG, and RMSE. Annoy performed better than other models.

**Dialogue-based Interactive Facial Editing via Reinforcement Learning**

- Developed a dialogue-based facial editing system via reinforcement learning that enables the system to explore long-term rewards that help with coarse-to-fine facial editing and a user simulator to provide feedback.
- The model was developed with Python using the **PyTorch** library and was able to achieve an accuracy of **80%**.

**Plasma Desk** with ExpressJs, NodeJs

- A web platform, driven by doctors, that connects an eligible plasma donor with a COVID'19 patient, based on compatibility. The details of the plasma donor are stored on the platform after clinical trials have been successful.
- The web portal was developed with Firebase as its backend and served using ExpressJs.
- The project won the **Best Medical Hack** and **Wolfram Award for Top 30 hacks** in MHacks 13 Beta Hackathon 2020.

## EDUCATION

**New York University | New York, USA** **Sept 2021 - May 2023(Expected)**
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

## 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