DISHA LAMBA

+1 (929) 444-1550 | dl4747@stern.nyu.edu | https://dldisha.github.io/ | GitHub | LinkedIn

EDUCATION

New York University, NY, USA Sept 2021 - May 2023
Master of Science in Computer Engineering
Relevant Coursework*:* Machine Learning, Deep Learning, Statistics, Big Data, Decision Optimization Models and Data Analytics
Course Assistant: Deep Learning, Computer Networking

Guru Gobind Singh Indraprastha University, Delhi, India Aug 2016 - Sept 2020
Bachelor of Technology in Information Technology

TECHNICAL SKILLS

Programming Languages:	Python, R, NumPy, Pandas, Keras, Scikit-learn, Matplotlib, React, Flask, HTML/CSS
Machine Learning:	TensorFlow, NLTK, Regression, Classification, Clustering, Decision Trees, Neural Networks
Quantitative Analysis:	Risk Management, Forecasting, Statistical Analysis, Time-Series Analysis, Optimization Models
Databases and Analytics:	SQL, Postgres, Hadoop, Spark, Kafka, Tableau, PowerBI, MS Suite (Excel, PowerPoint)
Cloud Services and Tools:	Amazon Web Services (AWS), Terraform, Snowflake, ElasticSearch, Git, Docker, Jupyter

PROFESSIONAL EXPERIENCE

Quantum Risk Analytics **Dec 2023 - Present**
Machine Learning Developer Intern New York City, New York
- Developed and optimized a predictive ML model, achieving **90%** accuracy via data pre-processing, feature engineering techniques.
- Collaborated cross-functionally to produce insightful data visualizations using **Tableau** for presentation to senior leadership.

Subconscious AI **July 2023 - Present**
Research Intern New York City, New York
- Optimized **Python ML model** with experimental design and testing for **25%** performance boost, **30%** data quality improvement.
- Automated content filtering with **Hate-Speech API**, achieving **90%** moderation process improvement and **saving 200+ hours**.
- Streamlined project timelines and resource utilization in a **Git** and **Agile environment**, enhancing the development lifecycle.

Data Glacier **June 2022 - Aug 2022**
Data Science Intern Remote
- Developed a **semantic text-ranking search engine**, enhancing user experience and **45% improvement** in operational efficiency.
- Designed **ETL data pipelines** using **SQL, Pandas** for data cleaning, reducing analysis time by **30%** for **1M** unstructured dataset.
- Attained **85%** accuracy using **Python-based BERT and TF-IDF** ensemble model for effective natural language processing.
- Optimized data storage & processing capabilities with **AWS S3, Lambda**, achieving **25%** data access improvement & cost reduction.

Sapio Analytics **July 2020 - June 2021**
Machine Learning Engineer Mumbai, India
- Spearheaded the development of **job-search portal** for the **Government of India**, generating **1.2M job opportunities** nationwide.
- Achieved **85%** accuracy in user-job matching with **K-nearest neighbor algorithm** in **Python**, tailored for user location and skillset.
- Collaborated with **SDE team** to seamlessly integrate ML model with **ReactJS**, enhancing user experience and engagement by **25%**.
- Devised **A/B experiments** with Product and Engineering teams for **data-driven decision-making** and product testing.

PROJECTS AND PUBLICATIONS

Data Analytics - Portfolio Optimization (MS Excel, Tableau)
- Revamped **deterministic model** with **MS Excel Solver** for portfolio optimization, resulting in **18% decrease in projected costs**.
- Developed **Tableau dashboards** showcasing portfolio performance, risk metrics, KPIs for informed investment decision-making.

Big Data - Movie Recommendation System (Python, PySpark, SQL)
- Utilized **SQL, Pandas, and Seaborn** for EDA, extracting valuable insights from a large **27M MovieLens** dataset.
- Built a baseline recommender system using **collaborative filtering ALS algorithm in PySpark**. Compared its performance against **Python LightFM** and **Annoy (ANN)** models, achieving an accuracy of **92%** to deliver top 3 movie recommendations.

NLP - Integrated System for Occupational Category Classification of Resumes (Python, NLP) - Research Paper
- Developed **NLP model** on 0.4M **resume dataset** to automate resume screening and categorizing candidates into occupational fields.
- Achieved an accuracy of **84%**, co-authored the paper at the **International Conference on Innovative Computing**, May 2020.

ACHIEVEMENTS AND LEADERSHIP

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