

## Education

- **Indian Institute of Technology, Bombay** Mumbai, India  
*Bachelor of Technology in Electrical Engineering* 2022  
 - B.Tech Project: Few-Shot Learning for Medical Image Diagnosis - [Slides](#)

## Technical Skills

- **Programming Languages** : Python, SQL, MATLAB, JavaScript, HTML, and CSS
- **Machine Learning** : NumPy, Pandas, TensorFlow, PyTorch, Keras, Scikit-learn, and OpenCV
- **Big Data Skills** : Apache (Spark, Airflow, Superset), MongoDB, Hadoop, Zepplin, and GCP(BigQuery)
- **Key Courses** : Machine Learning, Image Processing, Probability and Random Processes, Neural Networks and Deep Learning, Structuring Machine Learning Projects, Practical Deep Learning\*

## Work Experience

- **Data Scientist, Numberz.AI** Aug'23-Jan'24  
*Financial Advisory Chatbot - AI-powered Retrieval Augmented Generation (RAG) application.* Bengaluru, India
  - Aggregated **100+** public listed companies market documents to establish a comprehensive knowledge database.
  - Developed a voice processing system with **<100 ms** latency using transformer-based model for **transcript generation**
  - Employed **Amazon textract** for text extraction from various file formats, **LangChain** SQL agent for CSV files, **large vision models (LVM)** for image summarization, ensuring a holistic approach to information retrieval.*Invoice Named Entity Recognition (NER) - Computer vision based application for efficient accounting and operations*
  - Utilized advanced **natural language processing (NLP)** techniques to optimize text detection algorithms, resulting in an **F1-score** of **0.92** for invoice processing through **optical character recognition (OCR)**.
- **Software Developer, Teachmint Technologies Pvt.Ltd** Jul'22-Aug'23  
*ETL-Driven Analytics Pipelines and Data Integration* Bengaluru, India
  - Integrated **5+** external data sources into analytics pipeline for better predictive modelling & operational efficiency
  - Drastically reduced processing time by **90%** through effective data collection from **diverse sources** and optimizing the decoding step bottleneck (**20+ mins to <2 mins**).
  - Successfully **deployed** Apache Airflow for the analytics team, enabling efficient execution of ad hoc tasks.*API Development and Monitoring*
  - Developed scalable **REST API** using **Flask** framework, facilitating flawless communication.
  - Leveraged **PySpark** for statistical analysis on a real-time dataset comprising **100k+** events. Defined metrics, including Requests Per Minute (RPM), Latency, and Failure Rate, to evaluate and optimize the **CI/CD** pipeline.*Dynamic Dashboard Development*
  - Constructed **hierarchical data structures** for sales, enabling development of dashboards on **Apache Superset**.
- **Data Scientist Intern, Confederation of Empowerment Initiatives (CEI)** Jun'21-Aug'21  
*Deep Learning-based IMDb Movie Rating Prediction* Mumbai, India
  - Enhanced learning features by adding **attention masks** derived from movie posters and image inputs
  - Attained high accuracy by implementing and comparing **Transformer architectures** such as **T-5, GPT-3** and **BERT**, to incorporate **attention-based** mechanisms for text inputs

## Key Research Projects

- **E-commerce Analytics and Predictive Modeling for Olist** | *Kaggle Competition* - [Code](#) Jul'23
  - Implemented predictive models, **reducing delivery time** errors by **25%**, using **XGBoost** model.
  - Utilized natural language processing (NLP) based **word cloud** to extract trends from customer reviews.
- **Image Enhancement Recommender** | *Image Processing* - [Report](#) Jan'22-Feb'22
  - Leveraged transfer learning with **ResNet50**, utilizing feature extraction and fine-tuning techniques.
  - Obtained an overall accuracy of **94.82%** using multi-task model's predicted metadata and popularity score.
- **Autonomous Driving - Car Detection** | *Convolutional Neural Networks* - [Code](#) Jan'21
  - Built car detection system using **YOLOv2** object detection algorithm to generate **bounding boxes** around it
  - Used **Score-thresholding, Non-max suppression** methods to select best bounding box and achieved **76%** accuracy
- **Sentiment Analysis on Quora Questions** | *Natural Language Processing* - [Code](#) Aug'21
  - Scrapped and parsed raw text data for **BERT** pre-trained dense neural network with Transformer architecture
  - Improved test accuracy to **95.6%** by fine-tuning the hyperparameters in **Keras** using **Adam optimizer**

## Extracurricular

- Recent webinars - **Exciting Trends in Machine Learning, Best practices for building LLM-based applications**
- **Summer of Science Mentor- Maths & Physics club, IIT Bombay** May'21-Jul'21
  - Guided 4 mentees about the basics of Machine Learning, Deep Learning and Neural Networks
- **Dakshana Selection Test- JEE Scholarship Test** Dec'19
  - Volunteered in a team of 4 and successfully conducted the Grand Dakshana Selection Test-2019
- Completed a year-long certified **Volleyball** training course under **NSO Sports** 2017
- Awarded with the **Tritiya Sopan** for working as a scout under **The Bharat Scouts and Guide** 2015