

Debojit Nath

debojit10.github.io | nathdebojit2000@gmail.com | linkedin.com/in/debojit-nath/ | kaggle.com/denojitnath

EDUCATION

Indian Institute of Technology Madras <i>Diploma in Data Science</i>	2021 - 2023
Techno India Institute of Technology <i>Bachelor of Computer Application, SGPA: 8.5</i>	2018 - 2021 Kolkata, West Bengal
Kendriya Vidyalaya Fort William <i>CBSE, GPA: (class 12th) 75%, (class 10th) 85%</i>	2009 - 2018 Kolkata, West Bengal

EXPERIENCE

Data Scientist <i>Shortlyst AI</i>	March 2022 - July 2023
<ul style="list-style-type: none">Conducted an assessment of the user's behavioral patterns to discern deficiencies within the systemDeveloped LinkedIn data scraping bots and stored data in MongoDB databaseOptimized queries and created indexes to achieve 2M profiles/day speedAutomated stats generation of bot performance for easy monitoring via SlackFine-tuned DistilBERT model to predict industry tags for companies lacking available industry tags	
Data Science intern <i>Shortlyst AI</i>	November 2022 - February 2023
<ul style="list-style-type: none">Conducted an assessment of the user's behavioral patterns to discern deficiencies within the system	
Data Science Intern <i>Empliance Information Services India LLP</i>	June 2021 - December 2021
<ul style="list-style-type: none">Developed scripts for data extraction process from more than 50 sources to create Structured Data for the legal team to use for Compliance Check using Selenium and BeautifulSoupAutomated the process of downloading files of clients from MCA.gov, from logging in to saving them locally in a desired structure while handling captcha checkingFine Tuned HuggingFace model to classify Indian articles for Adverse Media Check with an accuracy of 95% and deploying an interactable webapp using Flask	
Data Science Intern <i>Iha Consulting Services Pvt. Ltd</i>	February 2021 - May 2021
<ul style="list-style-type: none">Built a Masked Face Recognition model using TensorFlow Object detection API and then converting to TFLite for optimized performance which performed better than out of the box solutions given they failed to detect faces	

RELEVANT COURSEWORK

Deep Learning Specialization <i>Certificate Link</i>	November 2019 <i>Deeplearning.ai</i>
Machine Learning <i>Certificate Link</i>	August 2019 <i>Stanford Online</i>
Mathematics for Machine Learning: Linear Algebra <i>Certificate Link</i>	February 2020 <i>Imperial College London</i>

PROJECTS

Toxic Comment Classification Challenge
<ul style="list-style-type: none">Participated in the Toxic comment classification challenge on Kaggle and trained a multi-label classification model which scored 0.9845 on public leaderboard
Fashion Apparel classification
<ul style="list-style-type: none">A dataset with imbalanced classes was used to train a Vision Transformer model with re-sampling method to tackle the class imbalance problem and improved the accuracy of the Vision Transformer model from 66% to 69%

TECHNICAL SKILLS

Languages: Python, C++, C

Frameworks: TensorFlow, PyTorch, Fastai, Flask

Libraries: Scikit-learn, HuggingFace, Keras, OpenCV, NLTK, Detecto, Selenium, BeautifulSoup