Mainak Chatterjee

RESEARCH ASSISTANT, **IIT Kharagpur**

 7980045963 | CHATTERJEE01M@GMAIL.COM | MAINAK-CHATTERJEE

EDUCATION

Maulana Abul Kalam Institute of Technology   West Bengal, India

Bachelor of Technology in Electronics and Communications Engineering (Hons.)   *Nov 2020 – June 2024*

- **CGPA :** 8.72/10
- **Affiliations :** Core Member of Literary Club TINT(LITWIS), Core Member of Coding Club TINT

EXPERIENCE

Indian Institute of Technology Kharagpur   07/24 – Present

Research Assistant with Saptarshi Ghosh   *West Bengal, India*

- ∗ Working on Multilingual Legal Natural Language Processing by leveraging advanced Large Language Models to develop and optimize language-agnostic solutions for complex legal texts across diverse linguistic frameworks

Indian Institute of Technology Kharagpur   03/24 – 06/24

Research Intern with Debarati Sen   *West Bengal, India*

- ∗ Successfully developed and implemented a comprehensive digital communication system using neural networks to enhance BPSK and QPSK modulation techniques
- ∗ Created a neural network model for BPSK and QPSK Systems with 10 input neurons and one output neuron, achieving 99 percent accuracy post finetuning
- ∗ Developed neural network models, including Feed Forward Neural Network and Convolutional Neural Network (CNN), to classify the MNIST dataset
- ∗ Achieved 99.9 percent accuracy with a Feed Forward Neural Network with 3 hidden layers

Indian Institute of Technology (Indian School of Mines) Dhanbad   06/23 – 08/23

Summer Intern with Samrat Mukhopadhyay   *Jharkhand, India*

- ∗ Successfully developed and implemented offline and online algorithms such as IHT, HTP, and FTHTPL in MATLAB for project titled 'Advanced Greedy Algorithms for Sparsity-Aided Generalized Online Caching'
- ∗ Utilized these Algorithms to calculate Reward and Regret showcasing a deep understanding of algorithmic optimization techniques
- ∗ Optimized the implemented algorithms by fine-tuning various input parameters, resulting in highly efficient and effective results

Suvidha Foundation   10/21 – 11/21

Machine Learning Intern   *Remote / Hyderabad, India*

- ∗ Collected, cleaned and augmented and pre-processed 42000 out of 50000 datasets from Kaggle for training deep learning as well as Machine Learning models
- ∗ Developed Flower Classification Model for classifying 56 different types of flowers
- ∗ Compared the accuracy of the trained models to find out the best suitable model

FEATURED ACADEMIC PROJECTS AND COLLABORATIONS

Multilingual Question-Answering Dataset for Indian Legal Queries   07/24 – Present

w/ Saptarshi Ghosh   *IIT Kharagpur*

- ∗ Designed a **Multilingual Question-Answering (QA) Dataset** consisting of **245 questions** in JSON format, with answers in **Yes/No** and corresponding explanations
- ∗ Focused on **Indian legal contexts**, creating multilingual datasets in **Bengali**, **Hindi** and **English** for cross-lingual evaluation
- ∗ Implemented **LLM-based workflows** using **Groq API**, **Google Translate API**, and **BNLP toolkit**, integrating state-of-the-art models like **GPT-3.5-Turbo**, **GPT-4o-mini**, **LLama-3**, **Gemma2** and **Mixtral** for generating answers and explanations
- ∗ Performed **evaluation by cross-checking LLM-generated answers with Gold Standard answers**, identifying mismatches and accuracy gaps across languages

Modern Digital Communication with Neural Networks · 03/24 – 06/24
w/ Swastik Chakraborty,Debarati Sen · *IIT Kharagpur*

- ∗ Developed a comprehensive digital communication system using neural networks to enhance BPSK and QPSK modulation techniques
- ∗ Implemented a Two Tap Rayleigh Fading Channel with Normal, ZF, and MMSE precoders to simulate real-world signal degradation
- ∗ Developed a neural network model for BPSK with 10 input neurons and one output neuron, achieving **99** percent accuracy

Advanced Greedy Algorithms for Sparsity-Aided Generalized Online Caching · 06/23 – 08/23
w/ Samrat Mukhopadhyay · *IIT (ISM) Dhanbad*

- ∗ Developed and implemented online and offline algorithms (**IHT, HTP, FTHTPL**) in MATLAB for calculating **Reward** and **Regret**, contributing to an innovative approach in the field of sparsity-aided caching
- ∗ Optimized algorithm code by meticulously **fine-tuning input parameters**, resulting in highly efficient solutions that demonstrated a deep understanding of algorithmic optimization techniques
- ∗ Plotted and visualized algorithmic outputs, specifically Reward and Regret, providing clear insights into their behavior and effectiveness

Flower Classification Model · 10/21 – 11/21
Suvidha Foundation

- Trained Machine Learning model from the data extracted from several Deep Learning Algorithms for **Image Recognition** and **Classification**
- Used VGG16 **Convolutional Neural Network** model for feature extraction from the input image datasets
- Used 2 versions of MobileNet CNN architecture for Image Classification and compared their testing as well as validation accuracy
- Trained Random Forest Classifier machine learning model using the data extracted from VGG16 for making the **best possible decision** regarding correct classification
- Achieved **57.12** percent as the highest testing accuracy on the trained machine learning model and **80.89** percent as the validation along **79.2** percent testing accuracy on MobileNet Version 2 CNN architecture

ACHIEVEMENTS

- **Finalist of Smart India Hackathon 2023 organized by the Government of India**

FEATURED COURSEWORK

- **Mathematics:** MIT RES-6-012: Introduction to Probability, MIT OCW; MIT 18.06: Linear Algebra, MIT OCW; Numerical Methods (4th Sem., MAKAUT)

- **Machine Learning:** Machine Learning Specialization, Stanford University; Applied Optimization for Wireless, Machine Learning, Big-Data IIT-K (NPTEL); Deep Learning Fundamentals - Intro to Neural Networks, Deeplizard; Neural Networks, 3Blue1Brown; Artificial Intelligence (8th Sem., MAKAUT)

- **Natural Language Processing:** CS224N: Natural Language Processing with Deep Learning, Stanford University; Finetuning Large Language Models, OpenAI; ChatGPT Promt Engineering for Developers, OpenAI

- **Computer Vision:** CS294-158 SP24: Deep Unsupervised Learning, UC Berkeley

SKILLS AND INTERESTS

Programming Languages: Python, Java, C/C++, MATLAB, LaTex, bash
Developer Tools: Google Collab, Huggingface, Kaggle, Spider, Jupyter Notebook, Visual Studio Code, Google Translate API, Groq API, Git
Libraries: PyTorch, TensorFlow, keras, Scikit-learn, NumPy, pandas, transformers