Soumitra Das

(+91) 8670234735 Chennai, India
Email LinkedIn Github

EDUCATION

Institute	Degree	Year	GPA/Marks(%)
Indian Institute of Technology Madras	M.Tech Data Science and AI	2024-2026	-
Indian Institute of Technology Delhi	M.Sc Mathematics	2020 - 2022	6.88/10
Ramakrishna Mission Vidyamandira	B.Sc(Hons.) Mathematics	2017 - 2020	7.75/10
WBCHSE	Higher Secondary	2017	85.2 %
WBBSE	Secondary	2015	84.14 %

RELEVANT COURSES

- **Sequence Models** [by deeplearning.ai]
- **Convolutional Neural Network** [by deeplearning.ai]
- **Getting Started with AWS Machine Learning** [by AWS]
- Pandas [by Kaggle]
- Introduction to Machine Learning[by Kaggle]
- intermediate machine learning[by Kaggle]
- **HuggingFace NLP Course** [by Huggingface]
- Build Basic Generative Adversarial Networks (GANs)(ongoing) [by deeplearning.ai]
- Data Structures and Algorithms, Mathematical programming, Numerical Analysis, Linear algebras, Data Mining, Cryptography, Graph Theory

TECHNICAL SKILLS

- **NLP:** Transformers, Language Models(Bert, GPT), MinGPT, Name Entity Recognition, Machine Translation, Question Answering
- **Computer Vision:** Object Detection [with YOLO],semantic segmentation [with UNet], Face recognition, Neural Style Transfer
- Deep Learning, Machine Learning, Basic GANs
- **Machine Learning Tools/Libraries: Pytorch, transformers, datasets, tokenizers,** scikit-learn, OpenCV, Keras, tensorflow
- **Programming Languages: Python C++, C ,JavaScript, MySQL**, R, Octave
- Others: Git,Github,Basic Web development, linux (debian based)

MSC THESIS

- **Image Encryption using two square cipher associated with Arnold map [Supervisor -** Prof. Rajendra Kumar Sharma](Deployed a python project for image encryption based on this paper.)

PROJECTS

- **Minbert for multitasking**: Here I created a Minimum representation of Bert and finetuned it for sentiment classification on SST and CFIMDB dataset. In the second part, I extended and finetuned the pretrained Bert for multiple downstream task(Sentiment Analysis, Paraphrase Detection and Semantic Textual Similarity) on SST, Quora dataset and SemEval STS Benchmark datasets.
- **Neural Machine Translation (Chinese to English) with attention RNN:** Implemented some of the major portion of the transformer architecture and used that model for machine translation.
- Chegg Expert Extension: Designed a chrome extension for easy navigation and notification for subject matter experts of the educational platform Chegg.