Data Scientist / AI Engineer
SHORYA SHARMA

+91 9528957619 | shoryamathura22@gmail.com
LinkedIn
GitHub
Ahmedabad, Gujarat (India)

Skills

- Python
- Machine Learning & ML Pipelines
- Probability and Statistics
- Deep Learning & Computer Vision
- Data Annotation for Computer Vision
- Data Analytics & SQL

- Transformers Pre-Training and Fine Tuning
- LLM Pre-Training and Fine Tuning
- PEFT (LoRA, QLoRA, Quantization)
- RAG techniques
- LLMs(LLaMA, Mistral, Gemma, Bert, GPT,Phi etc)
- Data Wrangling (Data Cleaning and Preprocessing)

Projects

Text Generation Using LLM (LLaMA-2 LLM Fine Tuning)

Executed Text Generation project leveraging Large Language Models (LLM), specifically fine-tuning the **LLaMA-2** model for customized text generation tasks using Python. Implemented cutting-edge techniques including **Quantization, LoRA, and QLoRA** for complexity reduction and efficient resource utilization. Resulted in enhanced model performance and reduced computational overhead.

Text Summarization Using LLM (LLM Fine Tuning)

Text Summarization project employing Large Language Models (LLM), fine-tuning the **'facebook/bart-large-cnn'** model using AutoTokenizer and AutoModelForSeq2SeqLM with TensorFlow. Implemented cutting-edge techniques including **Quantization, LoRA, and QLoRA** for complexity reduction and efficient resource utilization. Resulted in enhanced model performance and reduced computational overhead.

Next Word/Sentence Predictor Using LSTM

Developed a Next Word/Sentence Predictor utilizing LSTM architecture in TensorFlow and Python, integrating advanced preprocessing techniques and dataset preparation. Implemented **Word2Vec embeddings** for enhanced model performance, contributing to the advancement of natural language processing applications.

Emotion Analysis Based on Text Using Bert

Conducted Emotion Analysis on Text leveraging BERT Transformer architecture with TensorFlow and Python, utilizing Hugging Face Transformers library. Led preprocessing efforts including cleaning, tokenization, and padding to optimize model performance. Executed pre-training and fine-tuning processes, contributing to the development of robust emotion detection models.

Facial Expression Recognition - Deep Learning (Deep Learning & Computer Vision)

Developed robust **real-time facial expression detection system** using CNN and OpenCV, achieving accurate recognition of diverse emotions. Leveraged CNN for hierarchical feature extraction and fine-tuned the model for optimal accuracy, while applying OpenCV for seamless real-time image processing with camera feeds.

Human Activity Recognition by Smartphone (Machine Learning)

Implemented Human Activity Recognition using smartphone data, initially analyzing a dataset with **563 features**. Utilized variance threshold analysis to **select the top 17.26% of features**, reducing the set to 100 through feature engineering. Employed the correlation matrix method to address feature correlation, enhancing data clarity. Achieved a remarkable model **accuracy of 97%**, indicating significant prediction performance improvement.

Work Experience

Relevant Experience:

Data Science Intern – TwiLearn Edtech Pvt. Ltd. – Remote	May 2023 – August 2023

Other Experience (IOT):

IOT and Testing Engineer – CIMCON Software India Pvt. Ltd. – Ahmedabad	March 2022 – Present
Manufacturing Engineer (SMT) – Rashmi Rare Earth – Noida	JUN 2021 – FEB 2022
Manufacturing Engineer (SMT) – Transsion India Pvt. Ltd. – Noida	DEC 2019 – MAY 2021
Engineer Trainee (SMT) – Napino Auto and Electronics Ltd. – Gurgaon	JAN 2019 – NOV 2019

Certification & Licenses

Machine Learning Specialization – DeepLearning.AI
Deep Learning Specialization – DeepLearning.AI
Generative AI (GenAI) with LLM – DeepLearning.AI

Education

PGD (DATA SCIENCE) – Symbiosis Institute – Pune	2023 – Present
Bachelor of Computer Applications – Jaipur National University – Jaipur	2019 – 2022
Diploma in Electronics and Communication –B.S.A College of Engineering– Mathura	2015 – 2018