

ROHIT TIWARI

📞 9871626616 ✉ knowrohit.07@gmail.com 🐦 [Twitter](#) 🏠 [Github](#)

EDUCATION

Amity University Noida

Bachelor of Technology in Computer Science

August '19 – June '23

Noida, Uttar Pradesh, India

EXPERIENCE

TartanHQ

July '22 – Present

Data Scientist

- Successfully deployed and fine-tuned Document extraction OCR model using LayoutLM v3
- Built deployed models for Payslip Forgery and BlurDetection using Tesseract, Huggingface and Spacy NLP.
- Analysed Data inflow of Payslips and their metadata using Apache Superset (MySQL) + worked on custom annotation tool for OCR.

Trill Marketplace

September '21 – March '22

MLOps Engineer

- Built RNN model for scriptwriting using Transformers, GPT-3 NEO.
- Worked on ML systems using PyTorch, TensorFlow and JAX in production to scrape and learn from the clustered data of street wear customers.
- Built multimodal projects from OCR and Nyströmformer to perform evaluations on multiple downstream tasks on the GLUE benchmark.

GooseAI

April '21 – July '21

Data Science - Intern

- Built Recurrent Neural Nets for NLP and experimental ML with physics such as Phononics, dimensionality reduction and federated learning.
- Improved Masked language modeling: which has to predict some tokens that are masked in the input using Tensorflow and MASK.

PROJECTS

SCRIPTWRITING USING RNN | TensorFlow, Keras, Numpy, Caffe | [Github](#)

- a recurrent neural network to write scripts.
- Detects user input texts and generates a new movie script.
- More training epochs = better text generation.

Bernard-The first sentient AI | Hugging Face, Transformers, GPT-2, PyTorch | [Github](#)

- Bernard combines a sophisticated neural network machine learning model and scripted dialogue content.
- Using transformers and Decoding methods, it mimics the talking style of the user.
- Bernard's heart lies a complex autoregressive language model GPT-2 that utilizes deep learning to produce human-like text.

BigGAN | CUDA, Torchvision, PIL, CLIP, HuggingFace | [GitHub](#)

- Text-to-image generation using GAN, CLIP
- The generation of images using both a point from latent space and image class information as input.
- Using torchvision.utils for image segmentation.
- CLIP is to pre-train a neural language model and an image classification model using image data extracted from the Internet with their respective captions.

TECHNICAL SKILLS

Languages: Python, C/C++, Swift, SQL

Technologies: Anaconda, Apache, NVIDIA Cuda, Jupyter, PostgreSQL, Shopify, WordPress, Linux

Developer Tools: Git, Docker, Microsoft Azure, VS Code, AWS Sagemaker, Superset, PyCharm, iTerm

Libraries: PyTorch, CUDA, BigSLEEP, Numpy, Pandas, JAX, Keras, Scikit-Learn

Software: Microsoft Azure, Anaconda, Neural Designer (XAI platform)

EXTRA-CURRICULAR

- Member of Amity Linux Assistance Sapience Club(ALiAS)
- Zonal-Level Football Champion, BBFS FC