

ROHIT TIWARI

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

EXPERIENCE

TartanHQ

July '22 – Present

Data Scientist

- Successfully deployed on AWS EC2 and fine-tuned Document extraction OCR model using LayoutLM v3
- Built deployed models for Document Classification, Payslip Forgery and BlurDetection using Tesseract, Torchvision, Detectron2 and Spacy NLP.
- Analysed Data inflow, timestamps 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 and ChatBot model for scriptwriting using Transformers, GPT-3 and GPT-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.

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 Chatbot | *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.

Product Recommendation Engine | *VGG16, AzureML, Open-CV, HuggingFace, BS4* | [GitHub](#)

- Dataset of 44k images scraped from E-commerce site.
- Feature extractions using Cosine similiarities + VGG Hidden layers
- Predictions using similarity scores + Azure image tags
- Ensembled learning using VGG16 in computer vision + NCF for combined predictions, later to be deployed on AWS EC2

NFT-GAN | *torchvision, numpy, PIL, tensorboard, scipy* | [GitHub](#)

- It was developed for the purpose of creating NFT avatar and collectible projects using Gmapping and Gsynthesis.
- Allows you to configure the image generation process in such a way that you have complete control over how rare each and every trait is
- Generate JSON metadata for your NFTs that are in compliance with OpenSea metadata requirements

TECHNICAL SKILLS

Languages: Python, C/C++ , SQL

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

Developer Tools: Git, Docker, Microsoft Azure ML , VS Code, Amazon S3, AWS Sagemaker, AWS EC2, Superset, PyCharm, Cloudflare, iTerm

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

Software: Microsoft Azure, Anaconda, Neural Designer, AWS, WeightsBiases

EXTRA-CURRICULAR

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